

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Contractualism and intergenerational justice

Cormack, David Manson

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Contractualism and Intergenerational Justice

David Manson Cormack

A thesis submitted for the degree of
Doctor of Philosophy

Department of Philosophy

King's College London

February 2019

Abstract

This thesis extends an approach to justice grounded in the idea of agreement between equals to the question of justice between generations. Taking as its starting point the claim that distributive principles should be justifiable to all who will be subject to them – in this case, those who comprise the present and future generations – it begins by identifying two issues which militate against the extension of this method to the intergenerational context: the non-identity problem and the mutuality problem. In order to overcome the non-identity problem I appeal to – and defend – the claim that it is possible to wrong future people without necessarily harming them: from this, it can be inferred that we have humanitarian obligations toward future generations, by virtue of their status as rights-bearers. I then begin my discussion of the mutuality problem with a critical assessment of the idea of a universal contract with transgenerational scope. Despite rejecting this approach, I suggest that it captures certain fundamental intuitions that we have toward future generations, which complement an analysis grounded in indirect reciprocity. In this way, the mutuality problem comes to occupy a position of centrality within the thesis. To resolve it, I expound a theory of intergenerational mutuality which, I argue, allows the intergenerational context to be analysed as a collective action problem – in accordance with the standard interpretation of the original position.

Acknowledgements

In the time it has taken to research and write this thesis I have incurred innumerable debts to very many people. I would, however, like to express my profound thanks to John William Devine, Sarah Fine, and Andrea Sangiovanni: without their encouragement, advice, and insight, this project would not have come to fruition.

I would also like to thank my close friends and immediate family, who have borne the stresses and strains of this project with patience and good humour. Their support throughout has been invaluable, and it is to them, collectively, that this work is dedicated.

Table of Contents

1. Contractualism and Intergenerational Justice	6
1.1 Preliminaries	8
1.2 The Contract Approach	20
1.3 The Non-Identity Problem	34
1.4 Just Savings and The Mutuality Problem	38
1.5 Conclusion	43

Part I: Non-Identity

2. The Non-Identity Problem	48
2.1 Wronging, Harming, and the Person-Affecting Principle	52
2.2 Wronging and Political Contractualism	57
2.3 The Young Mother Objection	62
2.4 The Actual Persons Restriction	65
2.5 Conclusion	73

Part II: Mutuality

3. The Universal Contract and Intergenerational Justice	76
3.1 The Present Time of Entry Interpretation and the Circumstances of Justice	78
3.2 The General Assembly Revisited	83
3.3 Equality and Intergenerational Justice	86
3.4 Community, Reciprocity, and Posterity	97

3.5 Three Problems of Intergenerational Reciprocity	99
3.6 Conclusion	102
4. The Intergenerational Social Contract	104
4.1 Reciprocity	107
4.2 The Problem of Intergenerational Justice	110
4.3 Indirect Reciprocity, Mutuality, and the Social Bases of Self-Respect	114
4.4 Mutuality and Stability	121
4.5 Indirect Reciprocity and the Problem of Scope	123
4.6 Intergenerational Mutuality and the Problem of Justification	125
4.7 Conclusion	127
5. Intergenerational Mutuality: Objections and Responses	129
5.1 Four Objections	130
5.2 Posthumous Interests and the Original Position	133
5.3 Posthumous Interests, the Just Savings Principle, and the Social Bases of Self-Respect	137
5.4 Historical Non-Compliance and Enforcement	140
5.5 The First Generation, Non-Compliance, and Reciprocity	144
5.6 Conclusion	147
Conclusion	149
Bibliography	153

1. Contractualism and Intergenerational Justice

Over the course of the past fifty years, the idea of distributive justice has come to occupy a position of pre-eminence within political philosophy. This, in large part, is a direct consequence of John Rawls's utilisation of the concept in his seminal 1971 work *A Theory of Justice*. For Rawls, justice is understood to be the 'first virtue of social institutions', and so his enquiry is directed toward specifying a set of distributive principles to regulate what he terms 'the basic structure of society' – that is to say, those major institutions which 'distribute fundamental rights and duties and determine the division of advantages from social cooperation.'¹ The success of *A Theory of Justice* therefore gave rise to a series of broader debates about the appropriate way of arranging these institutions. At the same time, the concept has an obvious resonance with many of the foremost issues within twentieth-century political discourse: discussions around minority rights, the distribution of income and wealth, or the expansion of social security programmes are, at heart, debates over how major social institutions should apportion the benefits and burdens of cooperation at the societal level. So much to say, then, that the idea of distributive justice can be readily applied to some of the most salient issues within contemporary politics – and that this goes some way to explaining the ongoing philosophical interest in the concept. Supposing that this is the case, though, it raises obvious questions about whether or not considerations of this nature are equally applicable to those problems which are unique to our present circumstances. Indeed, I would want to argue that some of the most exigent questions being asked of electorates and policymakers in the early twenty-first century are structurally dissimilar from those of recent history. This is because rather than involving our contemporaries, they turn upon the rights and interests of future

¹ John Rawls, *A Theory of Justice* Revised Edition (Cambridge, MA: Belknap, 1999) p. 6.

generations. The most obvious – and urgent – example of this concerns anthropogenic climate change, but wider questions of demography and resource conservation also enjoy increased significance. With this in mind, it is worth noting that most enquiries into the way that major institutions distribute the benefits and burdens of social cooperation tend to focus explicitly on relations between contemporaries. In light of issues such as climate change, we therefore need to consider whether the simple fact that the affected parties are spread out over time is significant from the standpoint of distributive justice. And if, as I believe, it is, we need to address the concomitant issue of how we might incorporate these intergenerational concerns into our deliberations.

This thesis attempts to extend an approach to justice grounded in the idea of agreement between equals to the question of justice between generations. To be more specific, it takes as its starting point the claim that distributive principles should be justifiable to all who will be subject to them – in this case, those who comprise the present and future generations. This is characteristic of a contractualist approach to questions of distributive justice. In seeking to apply this to the question of justice between generations, I identify two problems which militate against its extension to the intergenerational context: the non-identity problem and the mutuality problem. To summarise, the non-identity problem poses a general challenge to the idea that persons ought to be the fundamental units of moral concern. Among other things, it therefore casts doubt upon the idea that interpersonal justification can be used to specify principles of intergenerational justice. By way of contrast, the mutuality problem is specific to contractualism and concerns the role of mutually advantageous social cooperation within the theory as a whole. On a standard contractualist reading, the question of justice arises in response to the need to regulate the distribution of the benefits and burdens associated with this kind of cooperative venture. However, certain unavoidable features of the intergenerational

context – most notably, the fact that generations are spread out over time – seem to rule out interaction of this sort. This thesis is therefore concerned with providing a contractualist response to both of these problems. My overarching goal is to arrive at an account of our distributive obligations toward future generations which is grounded in a conception of indirect reciprocity – and informed by an underlying theory of intergenerational mutuality. To that end, this introductory chapter serves to elucidate these two challenges to the contractualist position in greater detail. Prior to that, though, we first need to attend to certain preliminary issues. Consequently, this chapter will begin by discussing three fundamental distinctions in how we might conceive of the problem of intergenerational justice. By separating the intergenerational context out in this way, I intend to delimit the scope of this enquiry whilst giving some indication of the goals and objectives of the project as a whole. Subsequently, I will explicate the contractualist approach, and provide a qualified endorsement of it. From there, it will be possible to examine the problems encountered when attempting to extend this approach to the question of justice between generations: the non-identity problem and the mutuality problem. Finally, I will provide an indication of how I intend to respond to these two challenges and give an outline of the structure of the thesis as a whole.

1.1 Preliminaries

A diverse range of separate – and, at times, unrelated – issues can be grouped together under the broader question of justice between generations. It would therefore be appropriate to begin by giving some indication of the scope and nature of this enquiry. With this in mind, it would be helpful to draw three fundamental distinctions between differing approaches to the subject of intergenerational justice. The first of these relates to the contemporaneity of the parties: are the generations in question entirely discrete, or will they coexist at some point in time? The

second concerns whether we conceive of our obligations as being owed to past or future generations. So, we might understand the question of intergenerational justice to be about our antecedents, and the extent to which we should be willing to memorialise their achievements or redress historical injustices that they were responsible for. Conversely, our deliberations may be directed toward the interests and prospects of future generations. Then, in addition, we can subdivide the issues that might arise in either context into those which affect coexistent generations and those where relevant parties are spread out over time. Finally, we can approach the subject of what is owed to future generations at either the global or domestic levels. At this juncture, it is worth emphasising that the principal focus of this enquiry will be on what is owed to non-contemporaneous future generations at the domestic level. In elaborating on these distinctions, I therefore hope to offer a justification for focusing on this aspect of the intergenerational realm. At the same time, though, examining these points of divergence will serve to bring certain unique characteristics of the intergenerational context into focus. In this way, these three initial distinctions underpin the wider project and prepare the ground for later sections of this chapter.

i. Contemporaneity

The distinction between discrete and contemporaneous intergenerational concerns is a direct consequence of generational overlap. Simply put, at any point in time members of at least three generations will coexist. This fact gives rise to what I term ‘soft’ problems of intergenerational justice. That is to say those disputes which obtain between contemporaries, but where membership of a specific generational cohort is understood to be a source of inequity. So, to give a contemporary example, most developed nations are currently having to shoulder specific burdens associated with the fact that that people are living longer, and that a particularly large

generational cohort – those born in the mid-twentieth century baby boom – are approaching and entering into old age. Consequently, societies are facing a significant decline in the size of their working-age populations whilst having to meet the costs associated with an ageing population: healthcare, pensions, and so on. All things being equal, then, earlier generations will be required to make greater sacrifices than their predecessors, and this may be a source of injustice. Similar considerations can be brought to bear on the problem of global warming when considered from the perspective of those younger generations whose life prospects will be affected by our actions in the present. As before, the fact of being born into an earlier generation gives us the ability to shape the opportunities and conditions that our proximate descendants will face. Notice, however, that potential injustices can arise in the opposite direction: if, for instance, later generations were to neglect the rights and interests of their retired predecessors, then it would clearly be inequitable. Indeed, an implicit assumption within the debate over how to finance the pensions and care of the baby boomer generation in old age is that later generations have justice-based obligations toward the elderly; the issue is how to reconcile these with historical contingencies which threaten to make them unduly onerous. Instead, what is at play in all of these examples – and what distinguishes them from the kind of disputes that might obtain between contemporaries at the intragenerational level – is that at a particular time each generation finds itself in a position of power over others, and this raises the prospect of that power being abused. This is evident in the case of global warming, where the present generation are tasked with making decisions which will have profound consequences for their successors. Similarly, though, the charge that younger generations are having unjust burdens visited on them by those approaching retirement depends, at least in part, on the claim that their predecessors failed to mitigate those costs – through, for instance, adopting higher rates of

saving.² The idea, then, is that this kind of relationship – in which one generation is able to exert a disproportionate level of influence over the prospects of another – is distinct from anything which obtains at the intragenerational level. The question of intergenerational justice may, therefore, arise in the context of relations between contemporaries belonging to different generational cohorts.

We can, however, conceive of intergenerational justice in terms of non-overlapping generations, where the relevant parties belong to entirely discrete cohorts. This has particular relevance to issues such as global warming, whose costs will also be borne by people who are as yet unborn. Equally, though, considerations of this nature may arise when contemplating our obligations and duties toward our deceased forebears. The first thing to say, then, is that this dimension of the intergenerational realm involves parties who, at the time of deliberation, do not exist in any corporeal form. Indeed, in its purest articulation, problems of this nature involve parties whose lifespans will never overlap – and so, the question of them coexisting at some point in time does not arise. In this sense, relations between non-contemporaneous generations would appear to be radically different from anything that obtains between concurrent generations. On the one hand, our ability to exercise power over earlier generations is limited in this context due to them being deceased: we can choose whether to honour or traduce their posthumous reputations, continue or abandon their projects, and so on, but we cannot affect their life prospects. On the other, our ability to exert influence over future generations is vast: not only can we affect their circumstances and life prospects, but we can also determine their composition – by affecting, through our actions and omissions, how many people are born and their specific identities. The power relations which characterise

². See, e.g., David Willetts, *The Pinch: How the baby boomers took their children's future – and why they should give it back* (London: Atlantic Books, 2010) p. 80.

interactions between non-contemporaneous generations therefore possess features which are dissimilar to those which obtain between contemporaries. Indeed, within the wider literature, relations between past, present, and future generations are generally understood to turn upon an unavoidable asymmetry.³ That is to say that the disparity in powers between ourselves, our forebears, and our successors is an inevitable feature of the relationship between discrete generations. So, our predecessors were able to exercise a degree of influence over our circumstances which is broadly comparable to the power that we have with respect to future generations. Similarly, our ability to affect the posthumous interests of prior generations is equivalent to the power that our descendants have over the projects and concerns that we will entrust to them. Consequently, relations between discrete generations will never be on an equivalent basis, since we cannot improve the prospects of earlier generations or do anything to further the posthumous interests of our successors. In this regard, there is an obvious contrast to be drawn between this state of affairs and the possibilities for interaction between contemporaneous generations, where like-for-like cooperation can occur. If, for instance, an earlier overlapping generation has behaved well toward their successors, than their descendants can make a return in kind – by, say, providing for them in their retirement; such methods of requital are simply not available to discrete generational cohorts. For reasons which will soon become clear, I refer to this as the ‘hard’ problem of intergenerational justice, in contrast to those ‘softer’ problems which arise between contemporaneous generations.

With these two aspects of the intergenerational context outlined, it is now possible to offer a brief justification for focusing on the problem of justice between non-overlapping generations. To summarise, the principal distinction between these two approaches to the question of

³ E.g., Jane English, ‘Justice Between Generations’, *Philosophical Studies* 31 (1977): 91-104 pp. 96-8; Brian Barry, ‘Justice Between Generations’ in *Liberty and Justice: Essays in Political Theory* 2 (Oxford: Clarendon Press, 1991) p. 244.

intergenerational justice and their intragenerational analogue relate to the temporal distance between the parties and the effect that this has on the way in which they interact. So, at the intragenerational level, the relevant parties find themselves in a broadly symmetrical relationship: if they choose to cooperate, then everyone will stand to gain at the same time and in the same way. By way of comparison, relations between non-contemporaneous generations are wholly asymmetrical: the fact of being spread out across time militates against any like-for-like cooperation and restricts the forms of interaction which are possible. Finally, the relationship between coexistent generations sits somewhere in the middle: although they can cooperate directly, their interactions are characterised by disparities in power, with exchanges taking place over time and involving qualitatively different goods – as in the case of older generations receiving benefits in old age in return for their historical exertions. We can, then, conceive of relations becoming increasingly asymmetrical as the parties in question become more remote, up until the point of their being entirely non-contemporaneous. The necessarily asymmetric relationship between past, present, and future generations is, therefore, the feature that distinguishes the intergenerational context from the perspective of distributive justice. The first reason for focusing on discrete generations, then, is to address this property of the intergenerational realm directly. Indeed, it is for this reason that I refer to the question of what is owed to non-contemporaneous generations as the ‘hard’ problem of intergenerational justice, as against those ‘soft’ problems which obtain between coexistent generations. Quite simply, the ‘hard’ problem takes those features which differentiate the question of justice between generations from that of justice between contemporaries and examines them in isolation. By concentrating on the radical asymmetry which characterises relations between non-overlapping generations, I therefore hope to arrive at an account which has applications to both aspects of the intergenerational realm. There is, however, a second reason for focusing on non-contemporaneous generations: the long-term implications of issues such as anthropogenic

climate change. To give an illustrative example, although a majority of the carbon dioxide being emitted today will be absorbed by the ocean within a matter of centuries, an appreciable quantity will remain in the atmosphere for thousands of years.⁴ Indeed, this dovetails with a wider concern around what Stephen Gardiner has referred to as intergenerational ‘time bombs’. That is to say, policies whose costs are only realised by our non-overlapping descendants.⁵ There is, then, an obvious need to extend any enquiry into the nature of justice between generations beyond the contemporaneous level. This, when combined with the explicatory potential of a comprehensive response to the ‘hard’ problem therefore provide the rationale for limiting the enquiry in this way.

ii. Direction

The second distinction follows directly from the first, and, in particular, the dissimilar powers that we have with respect to past and future generations: to the extent that we have obligations toward them, they are going to differ in content and hinge upon separate concerns. Our obligations to posterity, for instance, are predicated on our ability to affect their interests and prospects. Indeed, it is for this reason that issues such as climate change and resource depletion have gained such traction within contemporary political discourse: they turn upon our ability to bring about palpable change in the circumstances visited on future people. By way of contrast, our duties to prior generations are primarily concerned with the relationship between the present generation and its forebears. So, when confronted with a historical injustice, the present generation is unable to affect the welfare of their ancestors; instead, the object of their deliberations will be a symbolic act of restitution, designed to restore the posthumous

⁴. See, e.g., David Archer et al., ‘Atmospheric lifetime of fossil fuel carbon dioxide’, *Annual Review of Earth and Planetary Sciences* 37 (2009): 117-34.

⁵. Stephen M. Gardiner, ‘A Contract on Future Generations?’ in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (Oxford: Oxford University Press, 2009) pp. 103-4.

reputations of those afflicted. Similarly, when considering whether or not to continue the projects of our ancestors, or memorialise their achievements, we are attempting to determine the extent to which the interests and rights that they had in life endure into the present. At this point, it is worth acknowledging that the question of what might be owed to former generations is often at its most salient when it relates to an existing injustice – as in the case of the historical institution of slavery and its legacy of racial inequality. It is, however, entirely plausible that obligations of this nature could arise without any corresponding contemporary issues – as in the case of a putative responsibility to honour the achievements of our predecessors. We are, then, attending to separate concerns when addressing the question of what is owed to past and future generations. In the case of our predecessors, their being dead prevents us from being able to affect their life prospects; we may, however, be able to benefit them by furthering the projects and interests that they had in life. Conversely, our actions and omissions in the present will have profound consequences for the prospects of future generations. We can, therefore, say that although comparable concerns motivate the question of what is owed to past and future generations, different things are at stake in either: consequently, it is helpful to distinguish one from the other.

Having brought this further distinction to bear on the question of intergenerational justice, it is now possible to provide a rationale for focusing on what is owed to future generations. Prior to doing so, though, I would want to emphasise that it would be a mistake to conceive of these two issues as being wholly unconnected; indeed, in Chapters Four and Five, I will argue that there is an interdependence between our obligations to past and future generations. Instead, this distinction would be better thought of as a device which allows us to focus on those features of the intergenerational context which have particular relevance to this enquiry. To that end, addressing the question of what is owed to posterity has three distinct advantages over an

investigation into our obligations toward our predecessors, or a general one encompassing both past and future generations. Firstly, the idea of distributive justice is more readily applicable to the problem of justice between the present generation and its successors. It is, for instance, possible to arrive at an account of what is owed to posterity by conceiving of the problem in terms of benefits and burdens which need to be allocated in an equitable manner. By way of contrast, although there is a distributive component in our obligations toward earlier generations, this method of analysis cannot provide a comprehensive account of our duties to our forebears. To elaborate on this, consider the claim that the present generation should seek to make reparations to the descendants of slaves. In that instance, a historical injustice is understood to engender distributive obligations in the present. Notice, though, that these obligations arise from the demands of corrective or restorative justice, rather than an explicit concern with the distribution of benefits and burdens across a society. In other words, even though a policy of reparations would have distributive consequences, its stated aim would either be to rectify historical wrongs or to promote reconciliation and seek atonement for past injustices.⁶ So much to say, then, that a thorough account of our obligations to prior generations would have to extend beyond the confines of a distributive analysis; to the extent that this enquiry is concerned with distributive justice, this represents a reason to concentrate on relations between the present generation and its successors. The second rationale for focusing on future generations relates to the two problems that emerge when attempting to extend the contract approach to the intergenerational context: simply put, they are more acute when considered from the perspective of future generations. There are, for instance, disturbing implications when the non-identity problem is applied to our relationship with prior generations: most obviously, the inference that, by virtue of our having worthwhile lives, we

⁶ See, e.g., Roy L. Brooks, 'Getting Reparations for Slavery Right – A Response to Posner and Vermeule', *Notre Dame Law Review* 80(1) (2004): 251-88 pp. 254-5.

are beneficiaries of those historical wrongdoings carried by our forebears – without which, we would not exist. However, the intuitions at play in the non-identity cases are more apparent when considering what we owe to future generations: the key move in Parfit's argument against the person-affecting view is to demonstrate that it leads to perverse conclusions regarding our conduct toward posterity.⁷ By focusing on the question of what is owed to future generations, we can therefore subject the contractualist position to greater levels of scrutiny. The third and final reason for taking our obligations to succeeding generations as the subject of this enquiry follows directly from the second: in addition to posing a bigger conceptual problem to the contract approach, there is simply more at stake in the question of what we owe to posterity. In saying this, it should be noted that our relationship to past generations is of considerable political importance – not least in terms of how a society conceives of itself. However, our ability to determine the prospects and composition of future generations, in conjunction with the urgency of issues such as climate change, makes the question of what we owe to future generations a singularly arresting political and philosophical problem. Indeed, given the exigent need to incorporate the rights and interests of future people into contemporary policy decisions, this aspect of the intergenerational context has particular relevance to any theory of distributive justice. In summation, then, the question of what is owed to future generations raises issues pertaining to the distribution of benefits and burdens across a society and possesses sufficient gravity to necessitate consideration from the perspective of justice. More pointedly, it intensifies those problems encountered when attempting to extend the contract approach to the intergenerational context. Consequently, by focusing on relations between the present generation and their successors, we can scrutinise the coherence and cogency of contractualism as a response to the question of intergenerational justice.

⁷ Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1987) pp. 361-2.

iii) Institutional Scope

The third, and final distinction reflects the institutional scope of any enquiry, and whether or not the question of what is owed to future generations will be understood on a global or a domestic basis. To elaborate, we can approach issues such as climate change from the perspective of their global impact or from the standpoint of specific countries. The same is true within discussions of intergenerational justice. It may, for instance, be the case that we have special obligations toward our direct institutional successors. At the same time, though, many of the issues which motivate an interest in intergenerational justice are particularly exigent and acute when viewed at the international level. In addition to this, questions can be raised as to how one set of obligations might relate to one another: if, for instance, there were a conflict between our humanitarian obligations to the global poor and those which obtain within domestic institutions, then we need to some means of resolving it. At their heart, these discussions turn upon the question of how specific distributive obligations are triggered. If, for example, we were to hold that the full demands of justice were owed to people as such, then the distinction between global, domestic, intra-, and intergenerational justice would dissolve. By way of contrast, it may be the case that we have relatively strong distributive obligations at the global level and weaker ones in the intergenerational context – and vice versa. Consequently, any enquiry into justice between generations has to specify its institutional scope: is the resultant theory to be implemented at the global level, or are its findings limited to the domestic level?

With the distinction outlined, it is now possible to state that this enquiry will be limited to the domestic level for two methodological reasons. Firstly, this thesis represents an attempt to

arrive at a theory of intergenerational justice by building on the existing Rawlsian framework. There is, then, a strong reason to delimit the scope of this enquiry to the domestic context on the grounds that it is from this perspective that Rawls approaches the question of justice between generations.⁸ Secondly, there is an extensive and varied literature concerning the application of Rawlsian contractualism to the international-cum-global context: to attempt to incorporate this into an enquiry concerning justice between generations would do a disservice to the complexity and scope of that debate.⁹ Instead, my hope would be that elements of this thesis can be brought to bear on theories in that field to elucidate our obligations to future generations at the global level. The decision to focus on the domestic level therefore serves to restrict the scope of our enquiry in such a way as to allow for a more detailed analysis of the contract method within the intergenerational context.

To recap, then, limiting the scope of this enquiry to non-contemporaneous future generations serves two main purposes. Firstly, it allows us to focus on those features which ultimately distinguish the intergenerational setting from its intragenerational analogue – most obviously, the asymmetric relationship between past, present, and future generations. Secondly, it exacerbates those problems which militate against the extension of the contract method to the intergenerational context. By concentrating on the relationship between the present generation and its discrete successors, then, we can address those issues that problematise the intergenerational realm for theories of justice generally, and contractualism specifically. And finally, directing our enquiry toward the question of what is owed to future generations at the domestic level enables a more rigorous analysis of Rawlsian contractualism within the

⁸ Rawls, *Theory* pp. 251-2.

⁹ See, e.g., Michael Blake and Patrick Smith, 'International Distributive Justice', *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), ed. Edward N. Zalta, URL: <https://plato.stanford.edu/archives/spr2015/entries/international-justice>

intergenerational context. Having dealt with these preliminary matters, it is now possible to offer an explication of contractualism as an approach to the question of distributive justice.

1.2 The Contract Approach

During the second half of the twentieth century, a number of approaches to questions of moral and political justification began to coalesce around the notion of mutual agreement as the basis of our obligations toward others. By and large, these theories were envisaged as correctives to utilitarianism and the aggregative tendencies within it.¹⁰ In this context, the metaphor of contract is enlisted to describe an initial agreement between parties understood to be equal participants in some form of cooperative endeavour. The moral and political principles selected in this way therefore derive a great deal of intuitive appeal from their ability to gain assent in conditions of equality. Indeed, as Stephen Darwall notes, morality is 'deeply implicated in the very notion of agreement, and vice versa.'¹¹ Put another way, the very fact that rational parties would agree to the terms of cooperation specified by this method gives them appreciable moral standing. The modern contract tradition therefore attempts to specify moral or political principles on the grounds of their acceptability to rational parties in ideal circumstances. To give a proper account of this approach we will, then, need to examine the nature of this initial agreement along with the cooperative venture it seeks to regulate. Before that, though, a distinction needs to be drawn between the two main intellectual currents within the contemporary contract approach: *contractarianism* and *contractualism*.

¹⁰ Rawls, *Theory* p. 24; David Gauthier, *Practical Reasoning: The Structure and Foundations of Prudential and Moral Arguments and Their Exemplification in Discourse* (Oxford: Clarendon Press, 1963) p. 126.

¹¹ Stephen Darwall, 'Introduction' to *Contractarianism/Contractualism* (Oxford: Blackwell, 2003) p. 1.

i. Contractarianism and contractualism

The contract approach therefore seeks to justify a set of moral or political obligations on the grounds of their mutual acceptability to equal parties in idealised conditions. It is, however, possible to distinguish two radically different interpretations of what such an agreement would look like. On one account, known as *contractarianism* the parties to any agreement are taken to be fundamentally self-interested. Accordingly, they will agree to terms of cooperation which maximise their interests and reject those which do not. Within the classical social contract tradition, this line of thought is most closely associated with the work of Thomas Hobbes; in contemporary theory, it manifests in the work of David Gauthier.¹² This may be contrasted with *contractualism*, according to which the parties are committed to selecting principles which can be justified to every individual who will be subject to them. In this way, the rational self-interest of the parties to any initial agreement is tempered by notions of moral equality and mutual respect between persons. Historically, accounts of this type can be discerned in the works of Jean-Jacques Rousseau and Immanuel Kant; in modern political philosophy, it is most closely associated with the works of Rawls and T.M. Scanlon.¹³ On such an account, the metaphor of contract is used to substantiate a set of moral norms and ideals which are independent of the theory itself, such as what it means to treat persons as equals in a given context. Conversely, the contractarian approach seeks to ground moral and political norms in the idea of instrumentally rational agreement between cooperating parties. To elaborate on this distinction, consider Stephen Gardiner's characterisation of the contract method as drawing upon three subsidiary concepts: equality, cooperation, and agreement. For the contractarian, equality is a purely descriptive concept

¹² *Loc. cit.*; David Gauthier, *Morals by Agreement* (New York, NY: Oxford University Press, 1986) pp. 4-6

¹³ Darwall, 'Introduction' pp. 4-5; Rawls, *Theory* pp. 10-15; T.M. Scanlon, *What We Owe to Each Other* (Cambridge, MA.: Harvard University Press, 1998) pp. 149-58.

which serves to ensure that the parties to any agreement have broadly equivalent powers; cooperation occurs on a mutually advantageous basis; and finally, the agreement itself is intended to specify terms of interaction which the parties would rationally accept. A contractualist interpretation, by way of comparison, will conceive of equality in normative terms – so the parties to the agreement will be understood to have equal moral worth. Similarly, cooperation will be characterised by stronger forms of reciprocal interaction, in which ideas of fairness and equity will inform the collaborative enterprise. And finally, any agreement will serve to justify the principles in question to those affected by them.¹⁴ Of course, we should be careful not to overstate this distinction. After all, Rawls makes use of rational choice theory within his initial agreement to demonstrate the preferability of his chosen principles to mutually disinterested parties behind a veil of ignorance. Similarly, a number of theorists within the contractarian tradition would argue that rational agreement can engender impartial and equitable principles.¹⁵ The point, however, remains that there exist two fundamentally different ways of conceiving of the contract method within contemporary philosophical discourse.

For the purposes of our present enquiry, we will be focusing on contractualism, rather than contractarianism. In part, this stipulation reflects a general sense of unease about the implications of a moral and political theory grounded exclusively in mutual advantage. As Gauthier himself notes, such a theory cannot generate obligations to those who cannot further our interests, such as young children or disabled people. Instead, these groups are understood to 'fall beyond the pale' of any moral or political obligations.¹⁶ More significantly, though, such an approach would appear to be severely limited in its applications within the

¹⁴ Stephen M. Gardiner, 'A Contract on Future Generations' pp. 79-80.

¹⁵ See, e.g., Johanna Thoma, 'Bargaining and the Impartiality of the Social Contract' *Philosophical Studies* 172 (2015): 3335-3355.

¹⁶ Gauthier, *Morals By Agreement* p. 268.

intergenerational realm. After all, due to the asymmetry in power relations between discrete generational cohorts, relations between present and future generations are effectively analogous to those between the powerful and powerless in the intragenerational context. Admittedly, there may be some scope for limited obligations between overlapping generations, but our responsibilities to posterity would begin and end with only our immediate successors. The contractualist approach, by way of comparison, offers some hope of addressing the distributive conflicts which characterise the intergenerational level in an equitable fashion. For example, asymmetrical relations between past, present, and future generations do not *necessarily* militate against the extension of the theory to the intergenerational realm. Similarly, the contractualist method allows for our having general, humanitarian obligations to future people which can be supplemented – rather than replaced – by those obligations arrived at in any initial agreement. Even before we account for the appeal of the theory, then, there are reasons to think that contractualism will be more readily applicable to the intergenerational context.

In saying this, it is important to emphasise that the simple fact of generations being spread out across time raises critical issues for any variant of the contract approach. As we will now see, this method of specifying principles of distributive justice presupposes some form of cooperation. The thought, then, is that obligations of justice are owed to those with whom we interact in a certain way; outside of this context, our obligations will be of a more general, humanitarian nature.

ii. Social cooperation

We now, therefore, turn to look at social cooperation within the context of a contractualist approach. As has been established, the idea of cooperation is a core feature of any contract theory. The idea, in outline, is that there are circumstances in which individuals will have to pool their talents to make themselves as well off as possible. In this sense, the parties to any agreement are understood to have a collective interest in cooperating to achieve this optimal outcome. At the same time, though, the members of this group will have heterogeneous preferences, leading to competing demands on the goods and resources secured through collaboration. It is in this context that the question of distributive justice arises. Put simply, the benefits and burdens of social cooperation need to be apportioned and, on a contractualist account, this needs to be regulated by principles that rational parties would willingly endorse. In this respect, Rawls's discussion of the basic structure of society is an obvious point of reference. On this account, justice is understood to be the 'first virtue' of basic social institutions, such as the political constitution and the fundamental economic and social arrangements. These institutions share three features which ensures their primacy in discussions of justice: they are non-voluntary, have a profound impact on the life prospects of all citizens, and are necessitated by social cooperation.¹⁷ The idea of the basic structure is therefore closely associated with that of mutually advantageous social cooperation: the institutions which comprise it generate those collective goods which give rise to cooperation in the first instance. The rules which govern these practices will therefore determine the nature of social cooperation and the distribution of the associated benefits and burdens across a society. The question of distributive justice, then, arises from the need to regulate the institutional framework which enables cooperation at the societal level and the collective benefits which flow from it.

¹⁷ Rawls, *Theory* pp. 6-10.

There is, however, another salient feature of these collective benefits which requires further attention: they are vulnerable to individual acts of free-riding. To elaborate, owing to the collective nature of the benefits produced in this way, and the difficulties involved in attempting to exclude individuals from them, it will be possible to reap the benefits of social cooperation without fulfilling any of the concomitant obligations. For instance, the members of society will benefit from public goods such as national defence which are financed through general taxation. Yet, if an individual were to stop footing their fair share of the tax burden, they would still be able to access the benefit in question. There will, in other words, be strong incentives to renege on the terms of any agreement. This unavoidable state of affairs therefore gives rise to a collective action problem. Simply put, although cooperation would make us all better off, the potential for free-riding will militate against any agreement. Rawls therefore amends the motivation of the parties to any agreement to include their being *reasonable* as well as rational. That is to say, in addition to being minded to pursue their own interests, they share a desire to cooperate with their fellow citizens on mutually acceptable terms: they have, in Rawls's turn of phrase, a capacity for a 'sense of justice'.¹⁸ In this way, the parties to any agreement can be satisfied that their fellow contractors will be minded to cooperate and willing to abide by the terms of any agreement. To expand on this, it is helpful to consider a classic formulation of the collective action problem: the prisoner's dilemma. In a standard prisoner's dilemma, two parties need to cooperate to secure the best possible outcome but fail to do so due to the fear of defection on the part of the other participant. We therefore find ourselves in a situation where the parties would prefer a collective strategy of cooperation but opt for an individual one of non-cooperation. By stipulating that the parties to any agreement will be reasonable – as well as rational – we are thereby able to check those anxieties that give rise to the collective action problem, and allow the optimal, cooperative

¹⁸ *Ibid.* p. xii.

strategy, to be chosen. That, of course, is not to say that the parties will be unconcerned by the potential for free-riding in any arrangements – indeed, as we will see, any agreement will depend upon the enforcement of the norms agreed to. Instead, reasonableness prevents the collective action problem from arising by ensuring that the parties to any agreement share a disposition toward cooperation. In this way, it builds a presumption in favour of reciprocal interaction into the theory. Thus, social cooperation within the contractualist approach is directed toward the provision of a central class of collective benefits, and the parties to any agreement will be understood to be both rational and reasonable – thereby overcoming the collective action problem.

iii. The original position

With the nature of the parties to any agreement and the subject of their cooperation dealt with, we may now address the agreement itself. Before proceeding, though, it is worth stressing that in all but a handful of cases, this 'agreement' is a representational thought experiment, designed to bring clarity to our deliberations. In the context of a contractualist theory, this involves structuring the agreement in such a way as to ensure that the principles selected could secure the assent of those who would be subject to them. To this end, Rawls's agreement – known as the original position – takes place behind a veil of ignorance which deprives the parties of morally arbitrary facts about themselves. Instead, their knowledge is limited to general facts about society and certain basic interests that they will have – most obviously, in maximising their share of primary social goods, such as liberties and opportunities, income and wealth, and the social bases of self-respect.¹⁹ The contractors are

¹⁹ John Rawls, *Justice as Fairness: A Restatement* ed. Erin Kelly (Cambridge, MA: Harvard University Press, 2001) pp. 57-8.

therefore unaware of their class, race, gender, talents, abilities, and so on. This, Rawls suggests, will ensure impartiality within any agreement and result in the selection of a set of principles that all rational and reasonable individuals would agree to. For instance, if we are unaware of whether or not we will have religious convictions, then self-interest will lead us to select principles which are consistent with freedom of belief. In this way, the instrumental rationality of the parties behind the veil of ignorance can be used to specify equitable terms of social cooperation.²⁰ Put another way, if we are unaware of all morally arbitrary facts about ourselves and our society, then we will automatically select principles to which all could agree. In this way, Rawls's contract functions as a heuristic device for delineating terms of cooperation that could secure the assent of all participants.

Before moving on, it would be pertinent to say something about the content of the principles selected behind the veil of ignorance. Given our lack of knowledge, Rawls suggests that a rational strategy would entail our adopting the perspective of the worst-off and looking to maximise their share of the benefits of social cooperation. The idea behind this proposal is that rational parties would be relatively indifferent toward any losses incurred through following this strategy as compared to the potential gains that could be secured through avoiding the unacceptable outcomes that another rationale may lead to.²¹ The contention, then, is that even if a great number of people deprive themselves of various benefits through adopting such a strategy, the resultant settlement will, at the very least, prove acceptable. Consequently, Rawls is able to argue that his two principles of justice as fairness would be selected over their competitors in the original position, precisely because they prioritise the interests and prospects of the worst-off. To summarise, these two principles comprise three

²⁰ Rawls, *Theory* pp. 118-23.

²¹ *Ibid.* p. 134.

component parts: firstly, the equal liberty principle, which states that persons are entitled to ‘the most extensive scheme of basic liberties compatible with a similar scheme for others’; secondly, the principle of fair equality of opportunity, which ensures that offices and positions are open to all; and finally, the difference principle, which stipulates that the only justifiable inequalities are those which make the worst-off as well off as they could possibly be. These principles are then ranked in lexical order so that the fair opportunity principle has priority over the difference principle and the equal liberty principle takes precedence over both.²² The lexical ordering of the principles of justice as fairness is of significance because, as we will see, Rawls conceives of our obligations to future generations as a constraint on the application of the difference principle. For our immediate purposes, however, the noteworthy feature of this analysis is the claim that by structuring the initial agreement in a particular way, mutually advantageous cooperation can be reconciled with ideals of moral equality and fair reciprocal interaction. In this way, the contractualist method is able to distinguish itself from the contractarian alternative whilst retaining the core features of the wider social contract approach.

iv. Contractualism's Appeal

Having given a brief sketch of the theory, we can now consider the appeal of contractualism as a method of specifying principles of justice. To that end, I intend to focus primarily on the role of reciprocity in the contract approach. Before that, though, something needs to be said about ideal agreement as a means of specifying principles of justice. In this respect, it is helpful to start from the metaphor of contract itself, and its use as a paradigmatic instantiation of an obligation. On such an account, agreeing to the terms of a contract is understood to have a

²² *Ibid.* pp. 266-7.

binding effect on the parties to it; if they renege on those terms without a legitimate excuse, then their actions will warrant some form of censure. Put another way, if one of the parties were to breach the contract, then they would be blameworthy precisely because they had previously agreed to it. In the case of a contract, then, the fact that it has been agreed to places the parties to it under an obligation. Of course, unlike the classical social contract tradition, contractualism is not a theory of political obligation; instead, it uses the idea of agreement to specify principles of distributive justice. There is, however, a corollary between the notion of agreement as the basis of our contractual obligations and its role in the selection of principles: in both cases, having agreed to something is understood to have justificatory force. Contracts, for instance, are binding because we have agreed to them. Similarly, if we can demonstrate that certain distributive principles would be the object of an ideal agreement, then we can provide a justification of them. This has especial relevance when we consider the pervasive and non-voluntary nature of those institutions which comprise the basic structure. There is, therefore, a need to justify the principles which regulate the basic structure to those who will be subject to them. In addition to that, though, the idea that the principles which regulate the basic structure could secure universal assent can be understood to facilitate a specific kind of social cooperation. To elaborate on this, consider Scanlon's discussion of the contract approach as allowing for a particular kind of coexistent relationship between people:

The contractualist ideal of acting in accord with principles that others (similarly motivated) could not reasonably reject is meant to characterize the relation with others the value and appeal of which underlies our reasons to do what morality requires. This relation, much less personal than friendship, might be called a relation of mutual recognition. Standing in this relation to others is appealing in itself—worth seeking for its own sake. A moral person will refrain from lying to others, cheating, harming, or exploiting them, “because these things are wrong.” But for such a person

these requirements are not just formal imperatives; they are aspects of the positive value of a way of living with others.²³

At this early juncture, it is worth emphasising that Scanlonian contractualism differs from its Rawlsian analogue in two ways: firstly, it is concerned with moral reasoning rather than the major social and political institutions of a society; and secondly, it attempts to specify principles that no one could reasonably reject – rather than ones which everyone would agree to.²⁴ In this instance, though, the basic insight is applicable to the political contract envisaged by Rawls. By cooperating on terms which could be endorsed by everyone we enable the members of a political society to relate to one another as equals. In Scanlon's analysis, the idea of mutual recognition captures this sentiment; for our present purposes, though, it might be better to think in terms of solidarity or to conceive of citizens cooperating on the basis of fair reciprocity. This, I suspect, is what Rawls has in mind when he suggests that the 'public recognition' of his two principles of justice can increase the effectiveness of social cooperation by bolstering the bases of self-respect.²⁵ By specifying principles of justice in a way that could be justified to those who will be subject to them, we therefore ensure that cooperation is to the benefit of all and proceeds on an equitable basis. We might, then, say that on the one hand, the idea of an ideal agreement is appealing because it is able to meet the justificatory challenge associated with the pervasive, non-voluntary institutions that comprise the basic structure. Equally, though, principles of justice which would be selected in these ideal circumstances embody a notion of fair reciprocity. In what follows, I am going to suggest that this underlying reciprocal logic is a decidedly appealing feature of the contractualist analysis, with potential applications in the intergenerational context.

²³ Scanlon, *What We Owe to Each Other* p. 162.

²⁴ *Ibid.* pp. 153-4.

²⁵ Rawls, *Theory* pp. 155-6.

Appeals to fair reciprocity can – and frequently do – form part of a wider egalitarian strategy, and the contract method can be understood to employ this notion in a particularly effective way. To give an idea of a general argument from reciprocity, consider the age-old claim that capitalism exploits workers and that this is unjust. On one reading, this injustice is understood to turn upon a lack of reciprocity between workers and the owners of capital, with the two groups receiving a disproportionate share of the benefits and burdens of social cooperation.²⁶ A related – albeit subtly different – line of reasoning can be used to inform a conception of disability rights: in that context, though, the complaint would be that existing social practices offend against fair reciprocity by preventing individuals from making a contribution in the first instance.²⁷ A generic argument from fair reciprocity, then, is likely to proceed from the claim that individuals or groups are either not receiving a fair return on their contribution, or are being excluded from the cooperative venture. In and of itself, then, the idea of reciprocity can be used to generate a powerful critique of the practices and institutions which obtain at the societal level. It is, then, significant that Rawls structures the original position in such a way as to 'embody the appropriate reciprocity and equality between persons [conceived of as free and equal].'²⁸ By embedding a conception of reciprocity in the initial agreement, Rawls therefore brings the contract method into alignment with a powerful line of social criticism. In this respect, those works which followed *A Theory of Justice* and used some variant of the contract method to scrutinise existing practices around social phenomena such as race and gender give some indication of contractualism's potential as a critical theory – understood in

²⁶ See, Paul Warren, 'Self-Ownership, Reciprocity, and Exploitation, or Why Marxists Shouldn't Be Afraid of Robert Nozick', *Canadian Journal of Philosophy* 24(1) (1994): 33-56 pp. 49-56.

²⁷ Steven R. Smith, 'The Social Construction of Talent: A Defence of Justice as Reciprocity', *Journal of Political Philosophy* 9(1) (2001): 19-37.

²⁸ Rawls, *Theory* p. 475.

the broadest sense of the term.²⁹ Although these authors interpret and deploy the contract approach in radically different ways – indeed, many have been sharply critical of aspects of Rawls’s methodology – some element of the reciprocal ideal plays a role in their respective arguments. At the very least, then, institutions are held to be unjust because they violate standards of reciprocity. At a more pernicious level, though, it may be that our existing practices militate against reciprocal interaction in the first instance, by inculcating attitudes of deference and dependence in specific groups and individuals.³⁰ In this regard, it is worth noting the evolution of Rawls’s own thought, as exemplified in his 2001 work *Justice as Fairness: A Restatement*. There, the difference principle is recast as a principle of reciprocity: the worst-off, then, are owed the assurance that they will be as well off as they could possibly be by virtue of their status as equal participants in socially cooperative practices.³¹ Indeed, to quote directly from Rawls:

The least advantaged are not, if all goes well, the unfortunate and unlucky—objects of our charity and compassion, much less our pity— but those to whom reciprocity is owed as a matter of political justice among those who are free and equal citizens along with everyone else. Although they control fewer resources, they are doing their full share on terms recognized by all as mutually advantageous and consistent with everyone’s self-respect.³²

²⁹ See, e.g., Susan Moller Okin, *Justice, Gender, and the Family* (New York, NY: Basic Books, 1989) pp. 89-109; Charles Mills, ‘Contract of Breach: Repairing the Racial Contract’ in Carol Pateman and Charles Mills, *Contract and Domination* (Cambridge: Polity Press, 2007) pp. 106-33.

³⁰ See Okin, ‘Justice, Gender, and the Family’ pp. 97-100. For an interpretation of Rawls as a critical theorist in the narrower sense of the term, see Amit Ron, ‘Rawls as a critical theorist: Reflective equilibrium after the ‘critical turn’, *Philosophy and Social Criticism* 32(2) (2006): 173-91.

³¹ Rawls, *JF* p. 49; pp. 123-6.

³² *Ibid.* p. 139.

This is indicative of a broader shift in Rawls's political thought, characterised by an increased scepticism toward the institutions of welfare state capitalism and greater emphasis on the ideal of 'property-owning democracy': an institutional arrangement in which the ownership of wealth and capital is dispersed across a society.³³ The capitalist welfare state, Rawls avers, fails to recognise a principle of reciprocity by allowing for high concentrations of property ownership.³⁴ Even within Rawls's articulation of the contract method, then, the idea of reciprocity is capable of generating a powerful critique of existing practices and institutions. This, I would suggest, is of relevance to our present enquiry for two reasons: firstly, the interactions between discrete generations can be analysed as a cooperative venture, with each cohort shouldering specific burdens and realising certain benefits. The idea of reciprocity may then have applications within the intergenerational context. Secondly, if, as I have suggested, a conception of reciprocity can motivate a wider critical project, then it would be well suited to the task at hand. After all, there can be little doubt that our existing institutions are inadequate with respect to the rights and interests of future people. By emphasising the role of reciprocity within the contract approach, I therefore want to suggest that it is capable of arriving at a plausible and attractive account of what is owed to future generations, whilst demonstrating the shortcomings in our current practices toward them.

Before proceeding to look at the non-identity problem, I would, however, want to make a final point about reciprocity and the contract approach. As may be evident, discussions of distributive justice take place at a certain level of abstraction, and this raises obvious questions about their practical value. It is, for instance, far from obvious why a problem like global warming needs to be analysed from this perspective. Indeed, there are numerous norms and

³³ *Loc. cit.*

³⁴ *Ibid.* p. 138.

values that can be appealed to in discussions of what might be owed to future generations – stewardship, sustainability, conservation, and so on – and it is understandable to question how a theory of fair shares fits into this conceptual landscape. In this regard, the idea of reciprocity can go some way to bridging the gap between theory and practice. Contractualism, as we have seen, starts from the idea of mutually advantageous social cooperation. The demands of justice, therefore, arise in response to the need to apportion the associated benefits and burdens in an equitable manner. In this way, the idea of distributive justice can be linked to a more concrete conception of relational equality – according to which, the goal of egalitarian justice consists in securing a particular kind of relationship between individuals to allow them to interact as equals.³⁵ A theory of fair shares, in other words, may be required in order to identify and prevent oppression and hierarchical social relations. There is, of course, considerable debate as to whether the interactions between generations can give rise to obligations of justice.³⁶ Supposing they do, though, then the asymmetrical power relations which obtain between generations would clearly warrant investigation from the perspective of justice. The intuition, then, is that a theory of distributive justice is required to ensure that the parties to any mutually advantageous social cooperation relate to one another as equals. By appealing to a conception of reciprocity, we may therefore be able to provide a theory of intergenerational justice with a concrete grounding in the relations that obtain between discrete generations at the societal level.

1.3 The Non-Identity Problem

³⁵ See, Samuel Scheffler, ‘What is Egalitarianism?’, *Philosophy and Public Affairs* 31(1) (2003): 5-39; Elizabeth Anderson, ‘What Is the Point of Equality’, *Ethics* 109(2) (1999): 287-337.

³⁶ See, e.g., David Heyd, ‘A Value or an Obligation? Rawls on Justice to Future Generations’ in *Intergenerational Justice* eds. Gosseries and Meyer pp. 167-188; Joseph Heath, ‘The Structure of Intergenerational Cooperation’, *Philosophy and Public Affairs* 41(1) (2013): 31-66.

Taking as its starting point the widely accepted claim that a necessary feature of our identity is our genetic endowment, the non-identity problem asks us to consider the myriad contingencies involved in one person coming into being. If, for instance, we were able to make slight alterations to the flow of history, then we could radically alter the composition of the current generation: had different gametes been involved in the relevant fertilisation events, then a different group of individuals would have been born. Thus, seemingly minor events can have profound consequences provided that they exert some influence over the particular sperm and egg which combine to form a zygote. In this sense, questions of existence are inextricably bound to the social and historical circumstances prior to conception. This fact gives rise to a serious problem for theories of intergenerational justice: if the current generation has a power to determine the composition of their successors, then we may be justified in imposing any manner of burdens on them, so long as their lives are preferable to non-existence. To expand upon this, consider Parfit's test case involving the depletion of resources, in which we are asked to choose between two policies: depletion and conservation. Choosing conservation will allow for the gradual increase in living standards, whilst electing to deplete resources will slightly enhance the prospects of those born in the next two centuries, whilst greatly depressing those of all subsequent generations. Clearly, depletion is intuitively the more problematic of the two policies. If, though, those generations born after the initial period of affluence owed their existence to the decision to deplete resources – as is highly probable – then our scope for objection would appear greatly diminished.³⁷ As against this view, Parfit suggests that we *can* reasonably prefer conservation to depletion, but only if we are prepared to conceive of morality in impersonal terms. Thus, instead of conceiving of our duties to posterity in *person-affecting* terms – where an act is wrong if and only if it makes a specific person worse off – we should reorient our

³⁷ Parfit, *Reasons and Persons* pp. 361-364.

deliberations around the effects of our choices on states of affairs.³⁸ Such a move, however, would conflict with the intuition that principles of distributive justice ought to regulate our conduct at the intergenerational level. At their most basic, impersonal approaches are characterised by an indifference toward the question of distributive justice: patterns of distribution matter only insofar as they bring about certain states of affairs. Indeed, more generally, it is difficult to see how an impersonal approach could accommodate concepts like individual rights and interpersonal obligations which necessarily apply to people. So much to say, then, that the move to an impersonal outlook is likely to have a transformative effect on how we conceive of our responsibilities to future generations. The non-identity problem therefore calls many of our considered judgements about justice and morality into question. It is, then, a problem that all theories of intergenerational justice need to attend to.

To give a better idea of what is at stake in the non-identity problem, it may be helpful to conceive of it as turning upon a logical inconsistency between three widely shared commitments:

- (1) Firstly, that an action is wrong only if it makes someone worse off. (Call this, ‘The person-affecting principle’.)
- (2) Secondly, that an act which brings someone into existence with a life that is worth living does not make that person worse off.
- (3) And thirdly, that the acts described in the non-identity cases – which bring people into existence with lives that are worth living – are, in fact, wrong.

³⁸ *Ibid.* p. 363.

There are, then, four possible lines of response to the non-identity problem: we either reject one of these three intuitions or amend them in such a way as to resolve the inconsistency. Parfit, as we have seen, favours rejecting (1) and adopting an impersonal approach in its place. We could, however, discard either the second or the third belief. Notice, though, that if we do, we find ourselves being forced to defend distinctly counterintuitive positions. In the case of rejecting (3), we have to countenance the possibility that the acts in question are wholly permissible; similarly, if we jettison (2), then we are required to make the argument that most lives are simply not worth having – and this may lead to the conclusion that procreation is wrong.³⁹ Consequently, the most popular alternative strategy to that suggested by Parfit has been to try and revise one of the three beliefs to make it consistent with the other two: in practice, the person-affecting principle is nearly always identified as the component of this antinomy which needs to be recast in some way.⁴⁰ In Part I of this thesis, I will look to build upon such an approach. Specifically, I will defend a line of response first put forward by Rahul Kumar which seeks to distinguish wronging from harming: the implication being that even if no one is made any worse off in a non-identity scenario, our actions can still be assessed as being unjust or blameworthy. This, I will argue, allows us to respond to the non-identity problem in a way that is consistent with the wider contract approach.

Although I have presented the non-identity problem as a general issue for all person-affecting approaches to questions of justice and morality – and this is how it tends to manifest in the

³⁹ See, e.g., David Heyd, *Genethics: Moral Issues in the Creation of People* (Berkeley, CA.: University of California Press, 1992) pp. 30-1; David Benatar, *Better Never to Have Been: The Harm of Coming Into Existence* (Oxford: Clarendon Press, 2006) pp. 40-8.

⁴⁰ See, e.g., Seana Shiffrin, 'Wrongful Life, Procreative Responsibility, and the Significance of Harm', *Legal Theory* 5 (1999): 117-48; Rahul Kumar, 'Who Can Be Wronged?', *Philosophy and Public Affairs* 31(2) (2003): 99-118; James Woodward, 'The Non-Identity Problem', *Ethics* 96(4) (1986): 804-31.

wider literature – I would argue that it raises two specific problems for the contractualist account. Firstly, as we have seen, the manner in which principles of justice are specified is necessarily person-affecting: the parties to the original position are representative of people who will have to abide by the terms of their agreement. It is, therefore, difficult to see how the contract method could be recast in impersonal terms without blunting its central justificatory strategy. Secondly, as Parfit notes, the parties to any agreement will automatically know that they exist at the point of deliberation. This, he argues, is analogous to assuming that we would all be men when choosing principles that would disadvantage women.⁴¹ There is, then, a question as to whether a contractualism can adjudicate on questions of existence in an impartial manner. The non-identity problem therefore calls key aspects of the contractualist analysis into question. To that extent, it represents a clear obstacle to the contract method's extension to the intergenerational context.

1.4 Just Savings and the Mutuality Problem

With the non-identity problem attended to, we can now turn to the second facet of the intergenerational context which raises problems for contractualism: the absence of mutuality between generations. In order to motivate this problem, it will first be necessary to say something about Rawls's proposed solution to the problem of justice between generations, in the form of the just savings principle. As has been established, the contract method attempts to justify principles by demonstrating that they would be the object of universal agreement in idealised circumstances. Consequently, any principles of intergenerational justice must be selected in this way. Rawls's formulation of the original position, however, problematises

⁴¹ Parfit, *Reasons and Persons* p. 392.

this task from the outset: under what is known as the *present time of entry interpretation*, the parties to his initial agreement know that they are contemporaries.

This, Rawls suggests, means that the parties to the original position have no reason to select any principle of intergenerational saving. As a result, some aspects of the original position will have to be modified to allow the contract method to be extended into the intergenerational context.⁴² The first such amendment, suggested in the original 1971 edition of *A Theory of Justice*, urged contractors to conceive of themselves as 'heads of families' possessed of a desire to further the welfare of their proximate descendants.⁴³ Aside from the problems such an assumption may raise with regard to the institution of the family, it also serves to limit the scope of our obligations to our closest relatives.⁴⁴ Consequently, in the revised edition of *A Theory of Justice*, this motivational assumption is allied to the observation that a fair rate of intergenerational savings can be specified by asking the parties to agree to a principle that they would want all prior generations to have followed.⁴⁵ By introducing this constraint into the structure of the original position, Rawls is able to motivate a concern for other generations among the parties to any agreement and argues that they would select a two-stage *just savings principle* to regulate our intergenerational conduct. This principle of intergenerational saving is directed toward establishing and maintaining just institutions over time. So, in the initial phase, the demands of justice require sufficient levels of capital to be transferred to future generations to allow for the realisation of just institutions. Once such an institutional framework is firmly established, however, net saving may fall to zero – effectively entitling subsequent generations to enough and as good as their predecessors.⁴⁶

⁴² Rawls, *Theory* pp. 254-5.

⁴³ John Rawls, *A Theory of Justice* (Cambridge, MA.: Harvard University Press, 1971) pp. 128-9.

⁴⁴ See, Okin, *Justice, Gender, and the Family* pp. 89-101.

⁴⁵ *Theory* p.111; pp. 254-5.

⁴⁶ *Ibid.* pp. 254-5.

Of course, for this line of reasoning to hold, we need to assume that all prior generations will have complied with their distributive obligations: if we suspect that there may have been historical non-compliance, then we would have no reason to agree to any savings principle. It is in this respect that the absence of mutuality between generations becomes especially problematic, because unlike at the intragenerational level where we can apply sanctions to enforce whatever norms are agreed upon, no such mechanisms are available. This is significant, because Rawls is clear that even if all citizens were moved by the same sense of justice such assurances would still be a necessary feature of any agreement.⁴⁷ In seeking to develop this argument, Heyd makes a distinction between full compliance and the enforcement of compliance. The former concept, he suggests, is only intelligible within the context of the latter. Thus, part of our rational basis for complying with those principles selected in the original position is the guarantee that others will be deterred from reneging on their commitments. Such deterrence, though, is impossible without mutuality.⁴⁸ The assumption that past generations will have complied with the savings rate chosen in the original position is therefore seemingly unwarranted. On the contrary, a lack of reciprocal interaction between generations would imply no saving whatsoever.

The lack of mutuality between generations therefore gives rise to two closely related issues for contractualist theories of intergenerational justice. On the one hand, the absence of like for like cooperation places future generations outside of any initial agreement. On the other, without some way of enforcing whatever principles are agreed upon, the parties to the original position have no reason to select a savings principle in the first place. It is for these reasons that many have identified an irresolvable tension between the idea of mutually

⁴⁷ Rawls, *Theory* p. 236.

⁴⁸ Heyd, 'A Value or an Obligation' p. 180.

advantageous social cooperation and the extension of the contract method to the intergenerational context. Heyd, for instance, argues that because we are unable to cooperate with our predecessors, non-compliance in any savings process can only be viewed as an act of disrespect – rather than an injustice – in contractualist terms.⁴⁹ Elsewhere, Daniel Attas is equally pessimistic, noting that the lack of mutuality between generations ‘poses an insurmountable problem to any...theory that aims to ground obligations on the ground of mutual advantage.’⁵⁰ Perhaps the most strident criticism of this nature, though, comes from Gardiner, who suggests that the absence of mutuality renders the contract method, as presently constituted, wholly inapplicable to the intergenerational context.⁵¹

The charge, then, is that Rawls’s justification of the just savings principle presupposes a form of mutually advantageous cooperation which is simply not possible in the intergenerational context: we cannot cooperate with our predecessors or our successors in a normatively significant way, and therefore the question of justice never arises. The mutuality problem therefore represents a challenge to the contractualist position to the extent that we believe that it needs to be extended to the intergenerational realm. Heyd, for instance, is content to fall back upon a set of general humanitarian obligations – akin to the duty of assistance that obtains between peoples at the international level – to regulate our intergenerational conduct.⁵² Gardiner, by way of comparison, speculates that ‘a refusal to countenance intergenerational issues may *undermine the appeal of the contract approach even for many of its core supporters.*’ The intergenerational context may, therefore, lead to a recasting of the contract approach as a theory of justice with universal applicability.⁵³ If, of course, the

⁴⁹ *Ibid.* p. 181.

⁵⁰ Daniel Attas, ‘A Transgenerational Difference Principle’ in *Intergenerational Justice* eds. Gosseries and Meyer p. 197.

⁵¹ Gardiner, ‘A Contract on Future Generations’ p. 86.

⁵² Heyd, ‘A Value or an Obligation’ pp. 183-7.

⁵³ Gardiner, ‘A Contract on Future Generations’ pp. 95-6; 114-6.

mutuality problem is predicated on a misconception and cooperation of the kind envisaged *does* obtain at the intergenerational level, then a contractualist theory of intergenerational justice, of the sort envisaged by Rawls would be perfectly intelligible. The mutuality problem can therefore be characterised in terms of a set of three inconsistent beliefs:

(1) Firstly, that those obligations associated with distributive justice arise within the context of mutually advantageous cooperation.

(2) Secondly, that principles of distributive justice are required to constrain intragenerational expenditure, with a view to securing the continuation of a just institutional framework.

(3) And finally, that the interaction between discrete cohorts at the intergenerational level cannot be characterised by mutuality.

To resolve the mutuality problem, then, one of these three beliefs needs to be rejected. On my interpretation, those who reject (1) are advocating some form of universal contract. The idea, in outline, is that by excising those commitments to mutually advantageous social cooperation from the theory, contractualism can be recast in such a way as to allow for its application to the intergenerational context. By way of comparison, those who reject (2) are going to accept that the obligations which regulate our intergenerational conduct are of a general, humanitarian variety. If this is the case, then Heyd's likening of the just savings principle to the duty of assistance between peoples at the global level would be appropriate: both reflect a broad commitment to justice as a value worthy of promotion, whilst stopping

well short of engendering obligations *of* justice.⁵⁴ Finally, those who reject (3) are able to incorporate the idea of obligations of intergenerational justice into the contract approach as it is presently constituted. Of course, the question facing such accounts is how mutuality might obtain between generations. In this regard, the obvious response is that generations cooperate *indirectly*. So, instead of collaborating on a like-for-like basis, discrete generations may be understood to cooperate by making a bequest to their successors which is broadly equivalent to the legacy they inherited from their forebears. Indeed, this dovetails with aspects of Rawls's account, particularly his characterisation of political society as a 'fair system of social cooperation over time, from one generation to the next.'⁵⁵ At the same time, though, it is not clear if cooperation of the form discussed could engender obligations of any kind – let alone obligations of justice. Brian Barry, for instance, likens the idea of indirect intergenerational reciprocity to being handed a toffee apple by a complete stranger: although we might enjoy it, and even feel a sense of gratitude, it is difficult to see how this engenders an *obligation*. Moreover, unless we can demonstrate that some form of mutuality obtains in the intergenerational context, then indirect intergenerational reciprocity of the kind envisaged will be irrelevant from the perspective of justice. Nevertheless, I would argue that this represents the most appealing line of response to the mutuality problem, and Part II of this thesis will be devoted to explicating a theory of intergenerational mutuality which engenders obligations of justice grounded in indirect reciprocity.

1.5 Conclusion

This introductory chapter has served three main purposes. Firstly, I have limited the scope of

⁵⁴ David Heyd, 'A Value or an Obligation' pp. 183-5.

⁵⁵ John Rawls, *Political Liberalism* (New York, NY.: Columbia University Press, 1993) p. 14.

this enquiry to focus on our distributive obligations toward non-overlapping future generations at the domestic level. Secondly, I have provided a brief outline of the contractualist approach – and suggested that the idea of reciprocity is an integral part of its appeal. And finally, I have introduced the non-identity and mutuality problems and illustrated why they militate against the extension of the contract method to the intergenerational context.

In the interests of expediency, I should state that this thesis will attend to the non-identity problem in Part I and the mutuality problem in Part II. So, Chapter 2 will argue that the non-identity problem can be resolved in manner that is consistent with the contractualist approach by distinguishing wronging from harming. It will also be suggested that we can use this framework to establish that we have humanitarian obligations toward future generations. From there, attention will turn to the mutuality problem as we move into Part II. This will begin with a critical assessment of the universal contract approach in Chapter 3. This will afford me the opportunity to demonstrate the preferability of an account grounded in a conception of reciprocity – and in doing so, foreground the mutuality problem within the structure of the thesis. To that end, I will look to refute the claim that the intergenerational context is devoid of mutuality in Chapter 4. The present generation, I will argue, depend upon future generations to preserve their projects and respect their posthumous interests. As against the claim that this lacks significance from the perspective of justice, I will demonstrate how these goods relate to the social bases of self-respect and the stability of a conception of justice. This, I will suggest, means that the parties to the original position would specify a principle of intergenerational savings on mutually disinterested grounds alone. Having expounded the theory of intergenerational mutuality, I will then address Brian

Barry's objection to indirect reciprocity as the basis of our intergenerational obligations. In Chapter 5, I will put forward four possible objections to intergenerational mutuality and respond to them in turn. In meeting these challenges, I will expand upon the central points of discussion in Chapter 4, bringing additional nuance and clarity to the argument. The idea of intergenerational mutuality is, then, the central concept in this thesis.

Finally, it would remiss not to take this opportunity to say something about the overarching aims and goals of this project. This thesis may, on the one hand, be interpreted as a relatively straightforward attempt to build upon Rawls's work in this area. At the same time, though, it can be read more specifically through the prism of my remarks on relational equality in this chapter. To elaborate, the Rawlsian contract method affords us an analytical framework which, as I pointed at the end of Section 2, is well suited to the task of identifying oppressive or hierarchal social relations. This has especial relevance to the question of what is owed to future generations due to the radical asymmetries in power which obtain in the intergenerational context: the simple fact that generations are able to act with impunity toward one another means that the powers at their disposal are liable to be abused. An unavoidable feature of the intergenerational context can therefore be understood to militate against the normative ideal that generations ought to interact with one another on equitable terms. Using the contract method to arrive at a theory of what is owed to future generations should, then, be seen as an attempt to specify what fair and equal terms of intergenerational cooperation would look like. By taking Rawls's theory as its starting point and emphasising the role of reciprocity within it, the thesis is therefore foregrounding the question of what it means to interact with future generations on an equitable basis. To that extent, the theory of intergenerational mutuality can be understood to bridge the gap between an abstract

conception of fair shares and the more concrete question of what it means to cooperate with our successors on terms characterised by mutual recognition and respect.

PART I

Non-Identity

2. The Non-Identity Problem

First outlined by Derek Parfit in his seminal 1984 work *Reasons and Persons*, the non-identity problem hinges upon the innumerable contingencies involved in a specific person coming into existence, and the ability of earlier generations to alter both the prospects and composition of future generations. Suppose, for instance, that we were to intentionally deplete resources to the detriment of our distant descendants. Instinctively, we might object to such an act by appealing to the deleterious consequences for future people. Such a response, however, overlooks the fact that those future generations will be comprised of people who owe their existence to our profligacy: had we chosen to conserve resources, then a very different set of events would have followed, in which different people would have had different children. Consequently, so long as those future people have lives which are worth living, the act of depleting resources in the present would appear to make no specific person any worse off. For those of us who understand people to be the fundamental units of concern within moral and political philosophy, the non-identity problem is therefore particularly troubling: it suggests that actions which are *prima facie* objectionable are, in actual fact, permissible. We might, then, characterise the non-identity problem as requiring us to resolve a logical inconsistency between three plausible commitments:

- (1) Firstly, that an action is wrong only if it makes someone worse off. (Call this, ‘The person-affecting principle’.)
- (2) Secondly, that an act which brings someone into existence with a life that is worth living does not make that person worse off.

(3) And thirdly, that the acts described in the non-identity cases – which bring people into existence with lives that are worth living – are, in fact, wrong.

In brief, Parfit avers that we ought to reject (1) and embrace an impersonal approach toward ethics, in which our deliberations are concerned with states of affairs rather than people. By way of contrast, David Benatar has suggested that being brought into existence visits innumerable unjustified harms on people, thereby invalidating (2).⁵⁶ Finally, David Heyd has suggested that any comparison between the welfare levels of existent and non-existent people is unintelligible, and that we should therefore be prepared to bite the bullet – and retain the person-affecting principle at the expense of (3).⁵⁷

As varied as these lines of response are, they all leave something to be desired. Moving away from the person-affecting view, as Parfit suggests, would require a thoroughgoing reappraisal of those fundamental moral concepts which we typically conceive of in personal terms – such as rights and obligations. At the same time, though, it is far from clear that our commitment to the person-affecting approach ought to be as unwavering as Heyd suggests: when choosing between our intuitive belief that the non-identity cases are wrong and a commitment to the person-affecting principle, Parfit's solution may be preferable. Finally, although almost all lives involve a degree of hardship, Benatar's argument requires us to accept that these harms cannot be outweighed by the benefits of existence: put another way, we have to accept the highly controversial conclusion that procreation is always and everywhere wrong. Consequently, I propose to resolve the non-identity problem by amending, rather than

⁵⁶ Benatar, *Better Never to Have Been* pp. 30-60.

⁵⁷ Heyd, *Genethics* pp. 93-126.

rejecting, the person-affecting principle. In its standard articulation, the person-affecting view holds that an action is wrong if it makes someone worse off. Since the non-identity cases involve people who otherwise not exist, they fall outside the scope of our ethical deliberations. That interpretation of the person-affecting principle, though, fails to capture our basic intuitions even in uncomplicated ethical scenarios. Suppose, for instance, that I promise to meet you at a certain time but arrive half an hour late without a good excuse. To make up for any inconvenience caused, I decide to buy you lunch – the benefits of which more than outweigh any of the losses caused by my being late. In such circumstances, my lack of punctuality would, in actual fact, have been to your advantage; it would, however, be strange to claim that that fact alone would absolve me of any blame. After all, in being late, I failed to live up to the promise that I made to you, and in that sense, I would remain culpable. Put another way, our actions and omissions have a moral significance of their own which cannot be captured by references to the positive or negative consequences which flow from them. In this respect, my analysis owes an obvious debt to the works of James Woodward and, in particular, Rahul Kumar, both of whom employed such a strategy to outline a non-consequentialist line of response to the non-identity problem. Following Kumar, I will look to draw a distinction between *wronging* and *harming*. To outline, a harmful action is a one which makes someone worse off than they would otherwise be; a wrongful action, however, is simply one which violates certain legitimate expectations that other people have with respect to our conduct. Consequently, the fact that future people cannot be harmed has no bearing on our deliberations: so long as they can be wronged, we can conclude that the non-identity cases are impermissible whilst retaining the person-affecting view. My analysis will diverge from Kumar's account in three ways: firstly, my interest in the non-identity problem concerns its implications for a theory of justice, rather than on broader questions within moral philosophy. Secondly, my approach is grounded in Rawlsian – rather than Scanlonian – contractualism.

And thirdly, I am going to argue that within the confines of a Rawlsian, justice-based approach we can limit the scope of deliberations to those future people who will exist, thereby excluding hypothetical people from our discussion. In this way, we can meet Parfit's objection to Kumar and respond to the non-identity problem within the confines of the contract method. Scanlonian contractualism and Kumar's analysis of the non-identity are, then, used to elucidate our response from the perspective of justice. By appreciating where their approach falls short, we can therefore illustrate the comparative resilience of the Rawlsian contract. My objectives in this chapter are twofold: firstly, I want to demonstrate that Kumar's proposed solution is both compatible with the political contractualism that forms the subject of this enquiry and more cogent when expressed in Rawlsian terms. Indeed, I want to suggest that by repurposing Kumar's analysis and applying it to the original position we can arrive at an account of our having humanitarian obligations to all future people who will exist. Secondly, I intend to demonstrate that the presumed consent of future people cannot serve as a justification for contemporary wrongdoing within the original position. This chapter can therefore be interpreted as an attempt to establish that we have general, rights-based obligations towards future people as such.

I will therefore begin by offering a brief sketch of the non-identity problem, along with Kumar's response to it, in Section 1. In Section 2, I will marry Kumar's analysis to the political contractualism which constitutes the subject of this enquiry – and suggest that it allows us to respond to the non-identity problem in a way that is both consistent with the wider contract approach and the claim that future generations will have rights. At that point, I will introduce a counterargument to Kumar's analysis in the form of the idea that identity-affecting rights violations might be justifiable on the grounds of the presumed consent of future generations – with specific reference to Parfit's examples of the young mother and the man who writes to

The Times. Section 4 will argue that from the perspective of justice this objection is dependent on one of two highly questionable assumptions: we either have to incorporate the notional rights of hypothetical people who will never exist into our deliberations; or, existence itself – rather than the goods which might be associated with it – is shouldering the justificatory burden in the argument. I will demonstrate that neither of these suppositions are tenable and affirm the rights-based approach in the process.

2.1 Wronging, Harming, and the Person-Affecting Principle

Let us begin by offering a more detailed explication of the non-identity problem. As has been established, the non-identity problem turns upon what Gregory Kavka referred to as the ‘precariousness of existence.’ Simply put, every person develops from a zygote comprising two specific gametes; had different sex cells been involved in the fertilisation event, then a different person would have been born. Consequently, everything from the identity of the parents to the timing of conception has a bearing on the composition of the next generation. To elaborate upon this, consider Kavka’s example of the slave child. In this scenario, we are asked to imagine a couple who enter into a contract with a wealthy individual requiring them to conceive a child who, upon being born, will be transferred into slavery.⁵⁸ Intuitively, the couple’s actions seem wholly immoral; on further inspection, though, it is difficult to see who is made any worse off by this transaction. After all, had the couple refused to sign the contract, then the child would not have been born. Assuming that the slave child has a life which is worth living, we are therefore forced to conclude that, in one sense, she is a beneficiary of this action. The issue, then, is whether an act can be wrong without being bad for a particular person. Building upon this, Parfit introduces a population-wide non-identity scenario in the form of

⁵⁸ Gregory Kavka, ‘The Paradox of Future Individuals’, *Philosophy and Public Affairs* 11 (2): 93-112 pp. 100-1.

depletion. In depletion, we are asked to choose between two policies: conservation and the eponymous depletion. Choosing conservation will allow for the gradual increase in living standards, whilst electing to deplete resources will slightly enhance the prospects of those born in the next two centuries, whilst greatly depressing those of all subsequent generations.⁵⁹ Due to the innumerable contingencies involved in their coming into existence, those people who come to inhabit those later generations will almost certainly owe their existence to the choice to deplete resources. As before, then, our intuitions are at variance with the person-affecting principle: so long as those populations in the depletion scenario have lives which they consider to be worthwhile, it is difficult to see how this act makes anyone worse off. From this, Parfit infers that any solution to the non-identity problem will need to appeal to impersonal reasons. That is to say, we will need to adopt an approach to ethics in which acts are right or wrong due to their bringing about a certain state of affairs, rather than their impact upon people. To that end, he contrasts the person-affecting principle with principle ‘Q’:

If in either of two possible outcomes the same number of people would ever live, it would be worse if those who live are worse off, or have a lower quality of life, than those who would have lived.⁶⁰

On such an account, then, bringing about a state of affairs in which the people who are born realise a lower quality of life than those who could have been born instead is taken to be objectionable – irrespective of the impact that this decision would have on specific people. In what are termed *same-people* choices – where the identities of the affected parties remain constant – Parfit suggests that Q is consistent with the person-affecting principle. However, in

⁵⁹ Parfit, *Reasons and Persons* pp. 361-2.

⁶⁰ *Ibid.* p. 360.

same-number choices – where different people will come into existence depending on the action that is taken – principle Q and the person-affecting view are rivalrous. Thus, in depletion, the person-affecting view would recommend that we deplete resources – since that would benefit the current generation and make no one any worse off. By way of contrast, principle Q would recommend conservation – since it would bring about a state of affairs in which the total quality of life was higher.⁶¹ Parfit, then, concludes that the non-identity problem cannot be resolved in a satisfactory manner unless we are willing to reject the person-affecting principle.⁶²

It may, however, be the case that Parfit's analysis of the person-affecting principle moves too fast and that, in actual fact, it can be articulated in a manner that renders it unproblematic. Recall, that Parfit understands the person-affecting principle as an approach to ethics which understands the rightness and wrongness of an act in terms of its effect on a specific person. Put another way, for something to be bad, it has to make someone worse off than they otherwise would have been. That, however, overlooks the possibility that some things might be intrinsically bad, irrespective of whatever consequences flow from them. Return, for instance, to the example of the slave child: we might reluctantly accept that, in one sense, she is the beneficiary of the contract, but still maintain that a grave injustice has been visited upon her. Consider, for instance, Woodward's case of a black man, Smith, who attempts to buy a plane ticket from an airline who refuses his custom on the grounds of his race. In the scenario, the flight that Smith would have been on crashes, killing everyone on board. Clearly, then, being the victim of racial discrimination has been to Smith's advantage. That, however, does little to change the fact that there is something intrinsically wrong in discriminating against people in

⁶¹ *Ibid.* pp. 368-70.

⁶² *Ibid.* p. 360.

this way.⁶³ Woodward suggests that this intrinsic badness can be captured by the idea of our having ‘relatively specific interests’ which can be violated without making us any worse off. We might, for instance, have such an interest in honouring our promises: even if I could greatly benefit you by reneging on a prior commitment, my actions would still be blameworthy.⁶⁴ The key idea, then, is that people have specific interests which are independent from considerations relating to their welfare. In this way, we can conceive of actions which might be bad for someone without making that person any worse off. The hope, then, is that by recasting the person-affecting principle to take account of the interests that people have, we may be able to resolve the non-identity problem without having to adopt an impersonal approach to ethics.

It is in this context that the distinction between wronging and harming becomes intelligible. As Kumar puts it, a central feature of all non-consequentialist ethical theories is that they understand that ‘*what one does* has an intrinsic significance in moral reasoning that is independent of *what happens* as a result of what one does.’⁶⁵ To elaborate upon this, consider the example that Kumar gives of the drunk driver. In the scenario, we are asked to envisage an evening walk which is briefly interrupted by an inebriated motorist swerving along the street: fortunately, no harm comes to us, and the incident is over so quickly that no distress is caused. The fact remains, though, that by choosing to drive in such a state, the motorist exposed any number of people to unjustified risks.⁶⁶ By failing to take account of the interests of others, we can, then, wrong someone without necessarily harming them. To develop this, Kumar introduces the notion of our having legitimate expectations with regard to one another’s conduct. These expectations are informed by a normative ideal of persons as ‘being capable of rational self-governance in the pursuit of a meaningful life.’ At the heart of this conception are

⁶³ Woodward, ‘The Non-Identity Problem’ pp. 810-1.

⁶⁴ *Ibid.* p. 810.

⁶⁵ Kumar, ‘Who Can Be Wronged?’ p. 105.

⁶⁶ *Ibid.* p. 103.

two separate claims about the nature of persons: firstly, they are understood to be capable of recognising and acting upon reasons; and secondly, they are able to reflect on these reasons and to select which of them to act upon. Consequently, we are able to appraise the actions, intentions, and beliefs of both ourselves and others in a manner that is independent of their consequences. In this way, legitimate expectations can apply to attitudes and motivations – in addition to actions. Indeed, as Kumar notes, we may be held culpable for ‘[failing] to have been responsive to certain considerations that it was legitimate to expect one to be responsive [to], or to have taken into account considerations...[that one ought to have] disregarded as irrelevant.’⁶⁷ Wrongdoing, in other words, can occur purely at the level of intention; unlike harm, it can be understood without reference to well-being. In this way, the person-affecting principle can be recast to reflect the legitimate expectations that we have with respect to one another’s conduct. A failure to meet these standards of behaviour can therefore be understood to violate fundamental commitments to both moral equality and mutual respect between persons. Following Kumar’s use of T.M. Scanlon’s contractualism in his explication of this idea, I will refer to this as the contractualist interpretation of the person-affecting principle: as we will see, by distinguishing wronging from harming, this approach provides the basis of a compelling response to the non-identity problem.

With the contractualist interpretation outlined, we can now consider how the legitimate expectations of future generations might be incorporated into our deliberations. To this end, Kumar appeals to a distinction between ‘type’ and ‘token’ identity statements. In this context, a token identity denotes a specific person or thing whereas type identity refers to the abstract category to which those things belong. Thus, a concept like ‘future generations’ could apply to the specific populations whose existence is predicated on our choices in the present or to a

⁶⁷ *Ibid.* pp. 106-7.

more general classification encompassing all potentially-existing individuals. Contractualism, as Kumar notes, is primarily concerned with determining ‘what a person, understood here as a *type* is entitled to expect of another person in certain *types of situation*.’⁶⁸ Thus, when attempting to select principles of intergenerational justice, our deliberations are directed toward the characteristics and interests that we can attribute to posterity as a whole, rather than specific future populations. We can, for instance, ascribe an interest in being treated as bearers of a capacity for ‘rational self-governance in pursuit of a meaningful life,’ to future people which, in turn, can be used to generate specific rights and entitlements.⁶⁹ In this way, Kumar suggests, a ‘discussion of the metaphysics of identity’ can be bypassed in favour of a ‘theory of responsibility.’⁷⁰ The idea, therefore, would be that actions like depleting resources could not be justified to posterity, and that this provides us with reasons to prefer conservation. The fact that token populations might benefit from the wrongful action cannot shoulder the justificatory burden. In this way, Kumar is able to argue that the non-identity problem need not arise in the contractualist framework: we don’t have to choose between recognising the wrongful nature of the non-identity cases and the person-affecting principle, because on this interpretation, they are convergent.

2.2 Wronging and Political Contractualism

With Kumar’s analysis outlined, we can now consider how the idea of types and wronging can inform the political contractualism which concerns us in this enquiry. In this way, we can begin to see how the concepts advanced in his account can inform on our theory of intergenerational justice. To begin, then, it might be helpful to restate – and expand on – the two challenges that

⁶⁸ *Ibid.* p. 111.

⁶⁹ *Loc. cit.*

⁷⁰ *Ibid.* p. 118.

the non-identity problem poses to contractualism as a method of specifying principles of justice. In this way, we can see how the appeal to types resolves the non-identity problem in a way that, at the very least, allows for general, humanitarian obligations to future generations. My aim in this section, then, is to demonstrate that contractualism can approach issues such as climate change from a minimal, rights-based perspective – of the kind outlined by Simon Caney.⁷¹

To recap, then, the non-identity problem raises two specific problems for contractualism as a method of specifying principles of justice. Firstly, the justifiability of principles on the grounds of their being the object of an ideal agreement is necessarily person-affecting: even if, for instance, we were to limit the principles on offer in the original position to impersonal ones, the parties to any agreement would still ultimately be representative of people. In order to recast the theory in impersonal terms, we would therefore have to abandon the idea that the principles of justice specified in this way could be justified to individuals. The move toward an impersonal approach would therefore have to come at the expense of an integral aspect of contractualism's appeal. Secondly, Parfit suggests that the contract approach is incapable of adjudicating on non-identity cases in an impartial manner. Recall that Rawls chooses not to make contractors ignorant as to their existence at the present time, fearing that such a move would detract from the plausibility of his thought experiment. Parfit concurs with this, observing that to do otherwise would require us to 'imagine something that we cannot possibly imagine.' Unlike Rawls, though, Parfit thinks that this introduces a bias into the selection of principles of intergenerational justice. This is because during the selection of principles, we will be aware of the fact that we will exist. For Parfit, this is analogous to assuming that we

⁷¹ See, Simon Caney, 'Climate Change, Human Rights, and Moral Thresholds', in *Climate Ethics: Essential Readings* eds. Stephen M. Gardiner, Simon Caney, Dale Jamieson, Henry Shue (New York, NY.: Oxford University Press, 2010) pp. 122-45.

would all be men when choosing principles that would disadvantage women.⁷² The charge, then, is that by limiting the scope of our deliberations to those people who will exist – and structuring any initial agreement accordingly – we violate standards of impartiality in the context of different-number problems. Insofar as contractualism derives a considerable degree of appeal from its claim to specify equitable terms of cooperation, this is a potential source of concern. In order to demonstrate how the appeal to types resolves these issues in favour of the contract approach, I will tackle these charges separately: firstly, I will employ Kumar's analysis to show how we might begin to mount a defence of the person-affecting view; and secondly, I will address Parfit's criticism with reference to the Rawlsian contract method.

Kumar's response to the non-identity problem therefore rests on an appeal to types – that is to say, the 'normatively significant cluster of relevant interests that a person might well have in virtue of certain general features of that person and circumstances in which she finds herself.'⁷³ This, by my reckoning, is entirely consistent with the contract method. Indeed, I would want to argue that the original position, as envisaged by Rawls, is best understood as an agreement between types of people, rather than token individuals. The veil of ignorance, for instance, is specifically designed to deprive the parties of any knowledge relating to the arbitrary features of their token identity. In turn, this ensures that principles of justice are selected on the basis of their justifiability to an entire society rather than their impact on specific people. To provide a useful analogy, Kumar likens the selection of principles in contractualism to a legal system: if we were legislating, then we would conduct our deliberations from the perspective of people could be referred to by the type description of 'citizens' or 'members of society', rather than

⁷² Parfit, *Reasons and Persons* pp. 391-2.

⁷³ Rahul Kumar, 'Wronging Future People: A Contractualist Proposal', in *Intergenerational Justice* eds. Gosseries and Meyer p. 265.

to particular token individuals who might satisfy that description.⁷⁴ Thus, as Elizabeth Finneron-Burns notes in her summary of Kumar's argument, 'if a person comes into existence with a cluster of interests characteristic of a standpoint' and we take that standpoint into account during our deliberations, then we have 'fulfilled the requirements of contractualist justification, regardless of whether or not it was also that principle that caused the person to exist.'⁷⁵ The original position, of course, is necessarily intragenerational in scope: due to the present time of entry interpretation, the parties to the agreement know they are contemporaries. What therefore matters, is that all things being equal, there will be future people, and that they will have rights.⁷⁶ Even if future generations are outside of the agreement envisaged by Rawls, we will therefore, have obligations toward them: hence, the just savings principle. As I have intimated, though, even in the absence of any institutional interaction between present and future people we would, on this account, still have humanitarian obligations toward them. We can, then, move from the claim that we can wrong future people to the more substantive one that they will have rights which institutions are obliged to respect. In this way, the non-identity problem can be resolved in a manner that is consistent with the wider contractualist approach and which can inform our existing practices toward future generations.

In responding to Parfit's claim that contractualism violates standards of impartiality, it would be expedient to quote directly from his discussion of the 'ideal contractualist' method:

⁷⁴ *Ibid.* p. 260.

⁷⁵ Elizabeth Finneron-Burns, 'Contractualism and the Non-Identity Problem', *Ethical Theory and Moral Practice* 19(5) (2016): 1151-63, p. 1155.

⁷⁶ See Annette Baier, 'The Rights of Past and Future Persons', in *Responsibilities to Future Generations. Environmental Ethics*, Ernest Partridge (ed.) (New York, NY.: Prometheus Books, 1981), 171-83; Joerg Chet-Tremmel, *A Theory of Intergenerational Justice* (London: Earthscan, 2009) pp. 48-9.

The principle we choose affects how many people exist. If we assume that we shall certainly exist whatever principle we choose, this is like assuming, when choosing a principle that would disadvantage women, that we shall certainly be men.⁷⁷

The contract method, then, is understood to be incapable of ensuring impartiality within our deliberations because the parties to any agreement will know that they exist. This, I would contend, is a distinctly unconvincing line of argument: it is not at all clear why existence should have any relevance from the perspective of the original position. Indeed, I would want to suggest that distributive obligations are owed only to those who do, have, or will exist. This is because having an interest in, say, maximising one's share of primary goods is dependent upon existence. Put another way, if we did not exist, then we would not have any interests in the first instance: in this sense, non-existence sets no one's interests back. Consequently, assuming the perspective of an existing person within our deliberations on the non-identity cases is not analogous to assuming a male perspective when adjudicating on matters relating to gender. Indeed, Parfit's analogy would only hold in a world populated exclusively by men, in which case assuming a male perspective would not introduce any biases into our deliberations. In this respect, it may be the case that Parfit is applying the contract method in an unorthodox manner. Suppose, for instance, that instead of attempting to justify a set of principles from the perspective of parties who are unaware of their circumstances, we instead assume the perspective of parties who *are* aware of their prospects. In that case, Parfit may have a point: if we approach the non-identity problem from the perspective of future people looking back at the historical contingencies leading up to their creation, then knowledge of our existence would prejudice our deliberations. However, that *ex post* method of justification runs contrary to the most fundamental insights within Rawlsian contractualism: principles of justice have to be

⁷⁷ Parfit, *Reasons and Persons* p. 392.

selected *ex ante* on the basis of general information about our life prospects, rather than particular knowledge of our specific circumstances. Consequently, I would argue that Parfit's appraisal of the contract method rests upon a conceptual misunderstanding of the Rawlsian original position. It should, however, be noted that the idea of existence justifying the wrongs involved in the non-identity cases is a recurrent one in criticisms of the contractualist approach.⁷⁸ I therefore want to use such an example to undermine Kumar's analysis. The idea, in outline, is that the retrospective consent of – wronged – future individuals can be used to justify the hardships inflicted on them. This, I will go on to argue, demonstrates the comparative desirability of the Rawlsian contract approach over its Scanlonian analogue in discussions of intergenerational justice.

2.3 The Young Mother Objection

The argument against Kumar's position that I have in mind suggests that, so long as people have lives which are worth living, the wrongs visited upon them could serve as a justification for our actions. To put this another way, future generations are unlikely to have legitimate expectations with respect to our present-day actions that would militate against their coming into being. In this regard, Parfit's example of the man born to a fourteen-year-old mother is particularly instructive. In it, a girl chooses to have a child at a particularly young age, giving her offspring a 'bad start in life.' As in any non-identity case, this seemingly obvious example of wrongdoing is problematised by the fact that the resultant child's identity depends on being born at the specific time; if their mother were to wait a number of years, then her child would be a different person.⁷⁹ Following Kumar, a contractualist response is likely to focus on the

⁷⁸ Corey Katz, 'Contractualism, Person-Affecting Wrongness, and the Non-Identity Problem', *Ethical Theory and Moral Practice* 21(1) (2017): 103-19, pp. 113-6.

⁷⁹ Parfit, *Reasons and Persons* p. 358.

legitimate expectations that types of children have with respect to their parents's conduct, and find that the young mother violated certain of these entitlements. Parfit, however, calls aspects of this approach into question when he scrutinises what is actually meant by a type identity description such as 'her child'. Statements of this kind, Parfit suggests, are only intelligible prior to the child coming into existence – when they could, conceivably, refer to something other than a specific person. At that juncture, it might make sense to claim that the girl's choice would set back the interests of 'her child.' Once the mother has the child, however, these words and phrases will refer to a specific person – who her actions will have benefitted. The distinction between types and tokens, then, is alleged to bring us no further forward in responding to the non-identity problem in a person-affecting manner. Indeed, as Parfit later puts it in *On What Matters*: '[g]eneral people are not individuals...[and]...sets of characteristics can't reject principles, nor could be owe anything to them.'⁸⁰ Kumar's analysis, then, can only be rendered intelligible if understood in impersonal terms: after all, that specific child is not being denied anything to which it would have been entitled, by virtue of the fact that it could not have existed in any other way. To expand upon this, consider those populations created through choosing a policy of depletion. Although it is evident that their existence would be the consequence of a historical wrongdoing, it is difficult to see how this relates to *their* legitimate expectations. Put another way, so long as we have lives that are worth living, it would be strange to invoke the expectations that attach to our type identity in order to bemoan the actions of our ancestors; in much the same way, the legitimate expectations of future generations are unlikely to act as an effective deterrent when the alternative would be non-existence.

⁸⁰ Derek Parfit, *On What Matters*, Vol. 2 (Oxford: Oxford University Press, 2011) p. 754 n. 236.

There is, however, another facet to the case of the young mother, which raises further problems for the contractualist account. In his discussion of rights and the non-identity problem, Parfit cites a letter sent to *The Times* by the – then adult – son of a teenage mother in response to a politician welcoming a decrease in teenage pregnancies. Although the man acknowledged the various hardships that had been visited on him by having such a young mother, he emphasised that his life was worthwhile. Indeed, not only did he not regret the choice his mother made, he was outraged that a politician should have implied that it would've been better for him not to be born.⁸¹ The benefits accrued through existence, then, serve to legitimate his mother's decision. We might, then, want to suggest that there are circumstances in which our legitimate expectations – and, in this case, our rights – *have* to be violated to serve some greater good: in this case, existing. To elaborate, consider Parfit's clarificatory example of a surgeon amputating the arm of an unconscious patient in order to save their life.⁸² In that instance, the patient's rights have almost certainly been violated, but the benefits – in this case, continued existence – justify the action. This is particularly problematic, because if the rights and legitimate expectations of future people can be overridden with the benefits of existence, then contractualism – at least of the form that I am looking to defend – cannot provide guidance in the non-identity cases. Indeed, if these reading is correct, then so long as future people have lives which are worth living, we are free to select whichever principles we want: their coming into existence will justify our impositions on them.

In summary, there are two ways of interpreting this line of argument: either in terms of presumed consent, or with reference to the idea that existence itself can justify our actions in the present. Returning to the example of the surgeon, we might want to say that her actions are

⁸¹ Parfit, *Reasons and Persons* p. 364.

⁸² Derek Parfit, 'Comments', *Ethics* 96(4) (1986): 832-72 p. 855

justified with reference to the thought that the patient would agree to the amputation if they were still conscious: we can, in other words, presume that they would consent, and this justifies violating their rights. We might, however, think that the benefits of continued existence are doing the work in this example: on this reading, we might acknowledge that a technical rights violation had occurred, but contend that this was justified by the gravity of the situation. In the case of a surgeon operating on an unconscious patient to save their life, both of these rationalisations strike me as being *prima facie* credible. However, I would want to argue that there is a difference between actions justified on the basis of continued existence – as in the case of emergency surgery – and actions which cause someone to exist. In the following section, I am going to argue that both of these justificatory strategies depend on an assumption which runs contrary to the person-affecting view as it manifests in Rawlsian contractualism: that the notional interests of hypothetical people – who will never exist – ought to play some role in our deliberations. My contention then, is that by limiting the scope of our concerns to those who will exist, the contractualist approach can refute the young mother objection.

2.4 The Actual Persons Restriction

Before offering a response to Parfit's example of the young mother, something ought to be said about the distinction between Rawlsian and Scanlonian contractualism. Up until this point, the two approaches have been understood to be broadly convergent. However, in light of the objections arising from the young mother example, attention must be drawn to the differences in justificatory methodology that sets one apart from the other. As, by now, we are well aware, the Rawlsian contract justifies principles on the basis of their being the object of agreement within the original position. By way of contrast, Scanlon understands a moral principle to be

justified so long as others could not reasonably reject it.⁸³ This is significant, because it means Scanlonian contractualism conceives of justification as involving actual people who can reject a moral principle. Consequently, when Parfit observes that general people cannot reject principles he highlights a fundamental tension between Kumar's discussion of types and tokens and the underlying precepts of the Scanlonian approach. Within the Rawlsian paradigm, however, there is no such contradiction: there, justification takes place at the level of the original position between representative parties. So much to say that, unlike Kumar, the Rawlsian contractualist can respond to Parfit with consummate ease: general people are not individuals, but they are *representative* of individuals and to that extent types can play a role in the justification of distributive principles. It is, then, at this juncture that we diverge from Kumar's analysis.

The case against the contractualist interpretation might be summarised as follows: to be properly person-affecting, legitimate expectations have to apply to token individuals of a specific type, and this opens the door to the possibility of token individuals either consenting to the wrongs being visited upon them or having their rights overridden by the benefits being conferred upon them. Put another way, if the populations in the depletion scenario knew that their existence was predicated on a violation of their legitimate expectations, then they would happily waive whatever entitlements they were due in order to secure their existence. Far from resolving the non-identity problem, then, we appear to be faced with the same dilemma that Parfit identifies: we can either accept that cases like depletion are no worse than conservation, or we can try to argue that these actions are wrong for impersonal reasons. In responding to this line of reasoning, I aim to show that it rests upon a controversial assumption which runs contrary to Rawlsian contractualism: either the presumed consent of future people or the

⁸³ Scanlon, 'What We Owe To Each Other' p. 85.

benefit of existence in and of itself can justify a wrongdoing. So, we are encouraged to imagine the future generations in the depletion scenario waiving whatever rights and entitlements they may be owed in exchange for their existence. That, however, supposes two things: firstly, that we can attribute interests to entirely hypothetical future people who will only come into existence if we act in a particular way; and secondly, that these notional individuals have an interest in existence for its own sake – that is to say, separate from any of the benefits that we might associate with it. As I will now argue, both of these claims rest on highly questionable foundations.

In the first instance, our attribution of interests to future generations is dependent upon their actually coming into existence: if, for whatever reason, there were no subsequent generations, then no individuals would exist to have those interests. For the same reason, it is difficult to see how anyone could have an interest in coming into existence. After all, had the young mother not had her child, then there would be no person whose interests could have been set back in this way. Of course, it is possible that, at some indeterminate point in the future, the young mother's offspring will look back on the circumstances of her birth and feel very grateful to have been born. Such a speculative exercise would, however, be impossible had she not been born in the first place. We cannot, in other words, seriously conceive of ourselves not existing. Consequently, an interest in existence is unlike any other interest we might attribute to people, because a failure to satisfy it makes the subject no worse off. Indeed, as both Finneron-Burns and Rivka Weinberg have noted, there is a kind of axiological neutrality to existence in and of itself: it is a prerequisite to someone having interests rather than an interest in itself.⁸⁴ In this way, an interest in existence is unlike any other interest we might attribute to people, because

⁸⁴ Finneron-Burns, 'Contractualism and the Non-Identity Problem' p. 1157; Rivka Weinberg, 'Identifying and Dissolving the Non-Identity Problem', *Philosophical Studies* 137 (2008): 3-18 pp. 13-4.

a failure to satisfy it makes no one any worse off. It is, then, difficult to see how it could shoulder the justificatory burden envisaged for it in the young mother objection: we can either visit a wrongful life on someone, or we can choose not to. If we elect not to bring that person into existence, then far from setting their interests back, we have prevented them from having interests at all. Consequently, when we discuss the legitimate expectations that a future generation has with respect to our conduct, we have to proceed on the assumption that the future generation in question will exist. For this reason, there is an obvious problem in attempting to compare the legitimate expectations of the populations in the depletion and conservation scenarios: only one population will come into being, and so only one will have interests which we can promote or hinder. Indeed, this feature of the contractualist interpretation serves to emphasise the importance of type identity to the wider theory: owing to the various contingencies involved, we cannot say whether or not a token individual or population will come into existence; we can, however, be quite sure in saying that types of people will exist in the future, and that they will have interests which we can either promote or impede. Thus, the idea that posterity's legitimate expectations could be outweighed by an interest in coming into existence is inherently problematic.

With respect to the argument from presumed consent, I would want to go further and suggest that it cannot shoulder the justificatory burden in the various non-identity cases. Indeed, in seeking to respond to the young mother objection, I would want to defend a stronger claim: that the presumed consent of a hypothetical person should be afforded no weight whatsoever in our deliberations. On first impressions, this may appear to be too forceful: after all, even if we face problems in attributing real interests to hypothetical people, we can at least conceive of people who would consent to our actions if given the chance. However, as Woodward observes, the notion of presumed consent can only hold water if we incorporate the consenting

agent into our moral community.⁸⁵ Consider, for instance, the invasive surgery example: although the surgeon violates certain rights to ensure that the person continues to exist, the presumably consenting party actually exists, and their wishes and interests are therefore relevant to our deliberations. Compare this with a scenario in which prospective parents can choose between two potential children prior to conception, one of whom will have a worthwhile existence, but one which – owing to some form of disadvantage – will fall below certain legitimate expectations that we might ordinarily attribute to future people. Now, the prospective parents may claim that the disadvantaged child would consent to whatever hardships were visited upon them to secure their existence. That proposed justification, however, depends upon our accepting that there will be a disadvantaged child who can reflect on the circumstances of their birth and retrospectively absolve their parents of any wrongdoing. We are, then, being asked to assume that that person will exist in order to justify whatever wrongs are inflicted upon them. That, however, is at variance with the spirit of the non-identity problem. After all, the problem only arises in the context of the contingencies involved in specific people coming into existence. We cannot, in other words, attribute the quality of existence to one set of individuals without prejudicing our deliberations. Consequently, we ought to exclude considerations relating to the presumed consent of specific future people from our analysis.

In turn, this focuses attention on the way in which Parfit formulates the non-identity problem. Recall, that on his account, cases such as depletion are wrong because they bring about a state of affairs in which a population realises lower standards of living than would have been the case had the other population come into existence. We might, however, wonder whether this comparison can actually be made. After all, we are comparing the real interests of actual people

⁸⁵ Woodward, 'Non-Identity Problem' p. 824.

with the notional ‘interests’ of hypothetical persons. To elaborate, imagine an argument for depletion which justifies our impositions on future generations with reference to their status as beneficiaries of the policy. So, without depletion they would not exist, and existence allows them to derive various benefits associated with a life worth living: therefore, the policy is in their interest. Now, Parfit would object to this line of argument by pointing out that although depletion might be good for those people, things could be much better in the conservation scenario. Such a comparison, however, rests upon a false equivalence between the existent and non-existent populations. To elaborate, imagine that we chose to conserve resources. Although our decision might confer certain benefits on those populations who come into existence and enjoy worthwhile lives, this does not come at the expense of the people who would have existed had we chosen depletion: they do not exist, and as such cannot benefit or lose out as a result of our decision. Put another way, not coming existence doesn’t deprive anyone of anything, because the person is notionally denied certain benefits never exists. There is, then, an obvious problem in attempting to compare the well-being and interests of future people who will actually exist with the rival claims of hypothetical people who can amount to nothing more than a figment of our imagination. Indeed, the only way to make such a contrast intelligible is to assume an impersonal metric: by, for example, replacing all talk of interests and expectations with units of happiness. In this regard, Parfit may be building a presumption in favour of his conclusion into the non-identity problem from the outset. In drawing attention to this, I should emphasise that although I have certain sympathies with Weinberg’s position, I am not attempting to argue that the entire non-identity problem arises as a result of a conceptual misunderstanding on Parfit’s part.⁸⁶ Instead, my remarks are better understood as proceeding along similar lines to those of Christine Korsgaard in her ‘Personal Identity and the Unity of Agency’. Namely, that there are underlying commitments in Parfit’s

⁸⁶ Weinberg, ‘Identifying and Dissolving...’ pp.

account which tend toward impersonalism in ethics and aggregative reasoning.⁸⁷ That is not to diminish from the quality his analysis, but it does allow us to better understand why, on such an account, Principle Q appears to be the appropriate response to the non-identity problem. On a contractualist interpretation, though, the only future people we have to justify our actions toward are those who will actually exist; hypothetical people are of no importance whatsoever. Consequently, our duties and obligations to posterity are necessarily limited to actual future people. This has the consequence of blocking the young mother objection, because the imagined wrongful act can only be justified if we afford the hypothetical person the status of someone who will actually exist.

Of course, it might be that Parfit has something else in mind when discussing cases such as the young mother. Weinberg, for instance, suggests that the issue at hand might relate to the fact that there will be actual future people who have been wronged, and that their having a worthwhile existence acts as a form of tacit consent to the historical wrongdoing.⁸⁸ The man who writes in to *The Times*, for instance, cannot condemn the actions of his mother because in giving birth to him, she granted him a worthwhile existence. Similarly, the people in the depletion scenario can have no regrets that their antecedents squandered so many resources, because it has enabled them to derive innumerable benefits from existence. The implication is, then, that the contractualist interpretation of the person-affecting principle cannot account for wrongdoing in real – rather than merely hypothetical – cases. The problem with this line of argument is that the move from not regretting one's existence to accepting all of the burdens associated with it is far from obvious. The people in the depletion scenario may, for instance,

⁸⁷ Christine M. Korsgaard, 'Personal Identity and the Unity of Agency: A Kantian Response to Parfit', *Philosophy and Public Affairs* 18(2) (1989): 101-32, p. 131; See also, Jeffrey Reiman, 'Being Fair to Future People: The Non-Identity Problem in the Original Position', *Philosophy and Public Affairs* 35(1) (2007): 71-92 p. 92.

⁸⁸ Weinberg, 'Identifying and Dissolving...' pp. 13-4.

acknowledge that their lives are worthwhile, but still feel that they have been subject to an injustice. Indeed, this brings us back to the earlier discussion about Rawlsian contractualism and *ex ante* justification. As was argued, principles of justice as selected on the basis of general facts rather than particular items of knowledge about our individual circumstances. Consequently, that there are particular people who, on Parfit's reading, have benefitted from having an inequity visited upon them is moot from the perspective of justice. Instead, what matters is that the ancestors of people in the depletion scenario acted on unjustifiable principles. In this respect, the slave child is a useful further example, because even though her existence depends upon her being enslaved, we can still clearly identify her as the subject of an egregious injustice. We might, then, respond to the man who wrote to *The Times* by acknowledging the contingent nature of his existence, and congratulating him on making the most of his life given in an inauspicious start, but nevertheless claim that, by virtue of his status as a person, he deserved a better start in life. Moreover, within the context of the contractualist approach, we can appeal to the idea of legitimate expectations to distinguish unreasonable grievances from complaints which are well grounded. Of course, the wronged party may wish to maintain that their appreciation of the benefits of existence more than compensates any hardships that have been visited upon them. To that, though, we can respond that not having been brought into existence would not have been bad for that person; indeed, as we have demonstrated, it would neither be beneficial or harmful. However, because that person *does* exist, they have rights and interests which entitle them to certain things: a childhood free from undue hardships, for instance. So much to say, that if we maintain that existence and non-existence cannot be compared in the way the non-identity problems supposes, then we can retain the idea that future people have rights and legitimate expectations which engender justice-based obligations toward them.

2.5 Conclusion

To conclude, it has been argued that non-identity cases such as depletion are unproblematic within the context of a Rawlsian contractualist interpretation of the person-affecting principle. This is because in addition to allowing us to draw an intelligible distinction between wronging and harming, it also allows us to exclude hypothetical people from our deliberations. In this way, we are able to address Parfit's example of the young mother in a way that other forms of contractualism – such as Kumar's – would find problematic. This is because within the Rawlsian approach, justification is done on the basis of general facts behind a veil of ignorance. Justification is not, in other words, owed to specific token individuals who are aware of their circumstances. In this way, it is possible to acknowledge the injustice of acts such as depletion even if they are responsible for bringing many people into existence who have lives which are worth living.

From the perspective of an enquiry into intergenerational justice, these findings may read like a formality. However, as I have argued, this line of response to the non-identity problem has profound implications for how conceive of our obligations toward posterity. The legitimate expectations that people have with regard to one another's conduct are necessarily varied – and depend upon how the relevant parties relate to one another. There must, however, be minimal levels of consideration and respect that people can insist on by virtue of their status as persons. This, I would contend, correlates with the idea of basic human rights. We can, therefore, be understood to have general humanitarian obligations to future generations. After all, those generations will be comprised of people, and they will have rights. This, I will accept, is but a starting point for a theory of intergenerational justice; it does, however, have practical implications from the institutional perspective. Obviously, I want to go further and suggest that

obligations of justice obtain at the intergenerational level, but this establishes a clear humanitarian baseline. Responding to the non-identity has, then, provided us with the foundations of a theory of intergenerational justice which I will expound in the subsequent chapters.

PART II

Mutuality

3. The Universal Contract and Intergenerational Justice

By virtue of the fact that generations are spread out across time they are able to exercise asymmetrical powers over one another. This has obvious implications for the contract approach, particularly in regard to the kinds of cooperation that can obtain at the intergenerational level. This may go some way to explaining why, in John Rawls's formulation of the original position – from which his two principles of justice are to be selected – the parties to the agreement are aware of their status as contemporaries under what is known as the *present time of entry interpretation*.⁸⁹ Our obligations to future generations within the orthodox contractualist framework are therefore limited. That, however, would jar with many of the core intuitions from which the contract method derives its appeal. After all, the original position is intended to overcome the disparities in bargaining power which characterise agreements in non-ideal conditions. The fact, then, that contractualism may be incapable of addressing a particularly severe – and politically exigent – asymmetry in power may well provide a decisive reason against adopting such an approach in the first place.

Suppose, however, that the contract method could be reimagined in such a way as to allow for an agreement between generations – either by altering the circumstances of justice or by doing away with them altogether. This, broadly speaking, is the ambition of a family of theories of intergenerational justice which can be subsumed under the heading of *universal contractualism*. Such an approach conceives of an initial agreement as involving every member of a political society across time. Consequently, the problems inherent in attempts to use the orthodox contract method to specify principles of intergenerational justice do not arise. Instead, we can move directly from the idea of mutually disinterested individuals selecting principles behind a veil of ignorance to a substantive theory of intergenerational justice. By dispensing with the present time of entry interpretation,

⁸⁹ Rawls, *Theory* p. 121.

therefore, the advocates of a universal contract hope to overcome the problem of justice between generations within the standard articulation of the contractualist method.

My reasons for dedicating a chapter to this approach are twofold: firstly, although it has a pedigree which can be traced back to Jane English's 'Justice Between Generations', it has, in recent years, emerged as one of the most promising methods of extending the contractualist approach to the intergenerational context.

⁹⁰ Secondly, by arguing against it, I can demonstrate the preferability of my favoured reciprocity-based account. This chapter will therefore argue against the idea of a universal contract and suggest that an amended version of the standard contractualist approach – complete with the present time of entry interpretation – can provide us with a stronger basis from which to defend and promote the interests of future generations. To do this, I will begin by offering a brief sketch of the circumstances of justice and the present time of entry interpretation in Section 1. In Section 2, we will consider the arguments for removing the present time of entry interpretation from our deliberations, along with Rawls's argument against doing so. It will ultimately be suggested that the major point of contention is moral rather than methodological – and that the central universalist claim is that the scope of justice *ought* to be extended to future people. With this in mind, Section 3 will attempt to reconstruct an argument for universalism on the grounds of moral arbitrariness, before moving on to suggest that this view depends on a thick notion of community. Having established this, Section 4 will then contrast the community-oriented approach with its reciprocal analogue and suggest that the latter is preferable to the former. Finally, in Section 5, I will outline three challenges to the reciprocity-based view and indicate the order in which I will attend to them in the following chapters. In this way, it will be argued that a theory of intergenerational justice which retains the reciprocal content associated

⁹⁰ See, e.g. Jane English, 'Justice Between Generations', *Philosophical Studies* 31 (1977): 91-104; Stephen M. Gardiner, 'A Contract on Future Generations' in Axel Gosseries and Lukas H. Meyer (eds.) *Intergenerational Justice* (New York, NY.: Oxford University Press, 2009) pp. 114-6; Merten Reglitz, 'The Practice-Independence of Intergenerational Justice', *Utilitas* 28(4) (2016): 415-40; Elizabeth Finneron-Burns, 'The Intergenerational Original Position', *Social Theory and Practice* 43(4) (2017): 805-23.

with the orthodox contract method represents a more propitious way of incorporating the interests of posterity into our deliberations. That, in turn, will serve to establish the mutuality problem as the principal obstacle to a contractualist theory of justice between generations.

3.1 The Present Time of Entry Interpretation and the Circumstances of Justice

Before going any further, it would be advisable to say something about the history and content of the circumstances of justice as an idea within political philosophy. In modern political theory, the idea that discussions of justice presuppose certain background conditions having been met originates in the work of David Hume – from whom Rawls borrows the phrase 'the circumstances of justice.' Although Hume's conception of the circumstances of justice differ from those employed in the contemporary contract tradition, there is a clear genealogical link between the former and the latter which is worth expanding on. On both accounts, questions of distributive justice are understood to arise as and when two or more people make competing claims on a certain good or resource within certain contexts. Firstly, conditions of moderate scarcity have to obtain: after all, if resources were abundant then the kind of disputes which give rise to the demand of justice would never occur.⁹¹ At the same time, though, it is suggested that under conditions of extreme scarcity the obligations of justice might be overridden by those 'stronger motives of necessity and self-preservation.'⁹² This latter stipulation represents an altogether more controversial aspect of Hume's analysis, since it can be read as implying that when our basic needs cannot be met, we can disregard any justice-based norms that might impede us. As we will see, though, Rawls's adaptation of the circumstances of justice mitigates certain of the more unsavoury conclusions that may follow from this observation. Of course, for disputes to arise, certain subjective criteria also need to be met: in short, there needs to be a heterogeneity of wants, and our capacity for altruism must, in some way, be limited. Indeed, as

⁹¹ David Hume, *An Enquiry Concerning the Principles of Morals* ed. J.B. Schneewind (Indianapolis, IN.: Hackett Publishing Company, 1983) p. 21; Rawls, *Theory* p. 110.

⁹² Hume, *Enquiry* pp. 22-3.

Hume points out, if everyone's interests were in harmony with one another, or our capacity for altruism was so developed as to allow for impartial reasoning, then we would have no need for distributive entitlements: there would be no point in '[raising] land-marks between my neighbours field and mine, when my heart has made no division between our interests.'⁹³ Thus, the potential for conflict is an essential feature of this account, and justice is taken to be the normative standard which is required to resolve these disputes in a fair and acceptable manner. At the same time, though, we should note that the circumstances of justice presuppose a conception of political society as a cooperative venture for mutual advantage. Indeed, Rawls draws our attention to this when he introduces the circumstances of justice as the 'conditions under which human cooperation is both possible and necessary.'⁹⁴ Additionally, then, the circumstances of justice have to take account of the various social and physical criteria which must be met if cooperative interaction of the kind envisaged is to obtain. It is in this context that the intergenerational dimension of our distributive obligations is so problematic. After all, the power that earlier generations wield over their descendants is, by definition, asymmetrical; posterity cannot interact with their forebears on equivalent terms. Having said this, disparities in bargaining power are also present at the intragenerational level – and one of the stated aims of Rawls's original position is to neutralise these differences by making the parties to any agreement ignorant of morally arbitrary facts about themselves and the society they inhabit. Much therefore hinges upon how the circumstances of justice are incorporated into the contractualist framework.

To appreciate the points of divergence between the Humean and Rawlsian accounts of the circumstances of justice, it is necessary to locate them in the wider context of their respective conceptions of justice. Rawls, for instance, conceives of justice as being 'the first virtue of social institutions' – the focal point of his enquiry being the selection of principles to regulate the manner

⁹³ *Ibid.* p. 22.

⁹⁴ Rawls, *Theory* p. 109.

by which the benefits and burdens of social cooperation are to be apportioned at the institutional level.⁹⁵ Hume, by way of contrast, orients his discussion of justice around a set of institutional conventions which arise over time, and which are justified by their utility – most obviously, those relating to property and contract.⁹⁶ There are, then, subtle differences in how the circumstances of justice are employed in either theory: on the Humean interpretation, they represent a series of conditions which have to be met before those norms and mores begin to develop; whereas on the Rawlsian account they refer instead to that set of features which make the relevant kinds of institutional interaction possible and necessary. When properly located within Rawls's wider theory, this seemingly minor amendment to Hume's initial understanding of the circumstances of justice has far-reaching implications. Simply put, it allows the discussion to be recast in terms of how an agreement might be reached with respect to the selection of principles. Hence, Rawls is able to use the circumstances of justice to inform his veil of ignorance. In this way, an initial agreement can be arrived at in conditions which ameliorate disparities in knowledge and bargaining power. Rawls therefore deprives the parties of particular facts about themselves and their society. At the same time, though, certain items of knowledge will be conducive to the kind of agreement that Rawls has in mind, and it is in this context that the parties to the original position are made aware of their status as contemporaries. Whether this is the case is a significant point of contention within the literature. For now, though, it is of note that Rawls's understanding of the circumstances of justice differ subtly from Hume's, and this allows him – and the contract method itself – to avoid the more obviously controversial aspects of the standard Humean account. The question, however, remains as to how such an approach can reconcile the idea of political society as a mutually advantageous cooperative venture with the unavoidable realities of the intergenerational context.⁹⁷

⁹⁵ *Ibid.* pp. 3-4.

⁹⁶ Hume, *Enquiry* pp. 27-30.

⁹⁷ For an explicitly constructivist reading, see Eric Brandstedt, 'The Circumstances of Intergenerational Justice', *Moral Philosophy and Politics* 2(1) (2015): 33-56.

Rawls, then, makes the parties to his original position aware of their status as contemporaries in order to bring about an agreement. We may, however, question why this is the case – particularly when we consider the problems that this move engenders. To elaborate upon this, it is worth contrasting the present time of entry interpretation with the idea of the original position as a general assembly involving every member of society over time. Were we to adopt the latter interpretation of the original position, then the principles of justice would, presumably, have intergenerational scope. Put another way, mutually disinterested parties who are ignorant as to which generation they will belong to will select principles which apply across time. The general assembly therefore represents an obvious way to attend to the problem of justice between generations. By way of contrast, the present time of entry interpretation appears ill-suited to the intergenerational context. After all, although the parties are ignorant as to the temporal location of their generation, the simple fact that they are aware of their contemporaneity could militate against their selecting any principles to govern intergenerational conduct. Indeed, as Rawls notes, on such an interpretation there is no reason for mutually disinterested parties to select a principle of intergenerational savings: prior generations will have either saved or not, and nothing they do can change that.⁹⁸ It may, therefore, seem odd that Rawls chooses the present time of entry interpretation over its competitor. To explain why, it is worth quoting Rawls's discussion of the general assembly interpretation at length:

[T]he original position is not to be thought of as a general assembly which includes at one moment everyone who will live at some time; or, much less, as an assembly of everyone who could live at some time...[i]f we conceived of the original position in either of these ways, the conception would cease to be a natural guide to intuition and would lack a clear sense. In any case, the original position must be interpreted so that one can at any time adopt its perspective. It must make no difference when one takes up this viewpoint, or who does so: the restrictions must be such that the same principles are always chosen.⁹⁹

⁹⁸ Rawls, *Theory* p.121; pp. 254-5.

⁹⁹ *Ibid.* p. 120.

The general assembly interpretation of the original position is therefore rejected on the grounds that it would 'cease to be a natural guide to intuition.' The implication, then, is that the present time of entry interpretation is needed to allow us to adopt the perspective of the original position. That, though, raises the question as to what differentiates generational membership from race, class, gender, and the myriad other facts about ourselves that the veil of ignorance blinds us to. The answer, in brief, is that Rawls thinks that knowledge of our existing at a specific time is essential from the perspective of the individuals imagining themselves in the original position. In this regard, we can imagine the conceptual difficulties in requiring someone who is aware of their existence at the present time being asked to imagine their being dead – as in the case of past generations – or unborn – in the case of future generations. Equally, it may be the case that Rawls fears that different principles would be chosen in the event of a general assembly: more will be said about this in Section 2, but it may be the case that certain forms of cooperation are only possible between contemporaries – and that this may be relevant from the perspective of justice. In either case, adopting the present time of entry interpretation leaves Rawls with the problem of attempting to extend his theory into the intergenerational realm, without undermining the bases of his theory. Initially, this took the form of an amendment to the motivational assumptions about the parties in the original position: in the original 1971 edition of *A Theory of Justice*, the parties are asked to conceive of themselves as 'heads of families' possessed of a desire to further the welfare of their proximate descendants.¹⁰⁰ Aside from the obvious problem such an assumption may raise with regard to the institution of the family, it also violates the notion that fair principles of justice are those which would be selected by mutually disinterested parties behind a veil of ignorance. Consequently, in his later writings on the subject, it is suggested that a principle of intergenerational savings can be generated by asking the parties to agree to a principle that they would want all prior generations to have followed.¹⁰¹ Although this seems intuitively plausible, it is not immediately clear *why* the parties to the original position would

¹⁰⁰ John Rawls, *A Theory of Justice* (Cambridge, MA.: Harvard University Press, 1971) pp. 128-9.

¹⁰¹ Rawls, *Theory* p.111; pp. 254-5; Rawls, *JF* p. 160.

entertain the question. So much to say that the present time of entry interpretation is inherently problematic from the perspective of a contractualist theory of intergenerational justice. It might, therefore, be reasonable to ask whether the general assembly interpretation of the original position might, on balance, be preferable.

3.2 The General Assembly Revisited

Indeed, given the problems occasioned by the present time of entry interpretation, it is worth considering whether the advantages outweigh the obvious drawbacks. Recall, that Rawls's principal reason for preferring the present time of entry over the general assembly interpretation relates to the intelligibility of the resultant thought experiment. In outline, he wants people to be able to picture themselves in the original position, and asking them to imagine themselves not existing at the present moment renders the deliberative exercise implausible. The idea, then, is that we can imagine ourselves as being born into a different socioeconomic class, or being assigned a different gender at birth, but we cannot credibly deny the fact that we exist in the here and now: it is a prerequisite of our being able to adopt the perspective of the original position at a given time. That, though, overlooks the insuperable difficulties already involved in assuming the perspective of the original position: we are asked to disabuse ourselves of myriad facts about ourselves, many of which are constitutive of our identities. For instance, as Jane English argues, the veil of ignorance already requires us to set aside fundamental aspects of our identity, such as our conception of the good; it is not obvious that this is any easier than denying our existence.¹⁰² The rigid dichotomy that Rawls delineates between existence on the one hand and contingent features of our identity on the other, is therefore highly questionable. More significantly, though, in his subsequent writing on the subject of the original position, Rawls appears to undermine this line of argument. As against the claim that the original position is metaphysically problematic, Rawls argues that it should be conceived of as a 'device of

¹⁰² English, 'Justice Between Generations' p. 99.

representation.' We can therefore sidestep any of the difficulties inherent in attempting to conceive of persons who are devoid of any of the constitutive parts of their identity; instead, the parties to the initial agreement are simply 'artificial creatures inhabiting [a] device of representation.'¹⁰³ The parties to the original position are therefore *representatives* rather than actual people. That, though, raises an obvious question as to why representatives need to be aware of their existence at a given time. Indeed, one can easily conceive of a general assembly in which mutually disinterested representatives from all generations select a set of principles with intergenerational scope. Rawls's principal argument in favour of the present time of entry interpretation is therefore unconvincing. Unless, then, we can find some additional reason against adopting it, the general assembly interpretation may be the most intuitively plausible way of extending the contract method into the intergenerational context.

There may, of course, be another reason why Rawls is reluctant to adopt the general assembly interpretation of the original position: namely, it might not yield the results he is looking for. After all, on the Rawlsian analysis, it is important that 'the restrictions must be such that the same principles are always chosen.' The implication, then, is that an interpretation of the original position which envisages the agreement as being one between non-contemporaneous representatives might select different principles of justice. On first impressions, this may seem odd: after all, the connection between the present time of entry interpretation and the selection of principles is far from obvious. Indeed, in her seminal analysis, English applies the two principles of justice as fairness directly to the problem of justice between generations in order to specify a fair rate of intergenerational saving.¹⁰⁴ On further reflection, though, there may be an issue with the difference principle within the intergenerational context. For instance, in his opening remarks on the question of justice between generations, Rawls notes that 'when the difference principle is applied to the question of saving over generations, it entails either no saving at all or not enough saving to improve social circumstances

¹⁰³ John Rawls, *Political Liberalism* (New York, NY.: Columbia University Press, 1993) pp. 27-8.

¹⁰⁴ English, 'Justice Between Generations' p. 92; pp. 100-3.

sufficiently so that all the equal liberties can be effectively exercised.'¹⁰⁵ The idea, then, is that conceiving of the original position as a general assembly introduces a degree of incongruity into our deliberations. Imagine, for instance, that we attempted to adopt the perspective of the intergenerationally worst off: would we look to maximise the prospects of the least advantaged group – at the possible expense of intergenerational savings – or look to secure the continued existence of just institutions over time? In this way, the present time of entry interpretation can be understood to simplify the deliberative exercise: by focusing on relations between contemporaries, the inevitable trade-offs that would otherwise characterise the selection of principles can be avoided. Of course, it may be the case that our intuitions can ultimately be clarified in the general assembly – perhaps, for instance, English is right that a favourable rate of savings can be specified by the difference principle. It is, however, the case that such an interpretation of the original position introduces an additional layer of complexity to the thought experiment. We may, therefore, surmise that the general assembly interpretation is rejected by Rawls due to its problematic nature within the structure of his argument from the original position: simply put, it dampens the intuitive appeal of the difference principle.

We should, however, acknowledge that there is a deeper, and more troubling respect in which the general assembly interpretation might be understood to militate against the selection of the difference principle. As Rawls notes in his subsequent writings, the difference principle is understood to reflect a deeper idea of reciprocity. The idea, in summary, is that by arranging the basic structure in such a way as to ensure that the least advantaged are as well off as they could possibly be, we ensure that the contingencies of birth work to the benefit of all in society.¹⁰⁶ Failing to do this, Rawls avers, means that those more fortunate individuals – who have been blessed with valuable talents, advantageous social backgrounds, and the like – are effectively using the worst off as a means to their own aggrandisement. Rawls's argument for the difference principle therefore hinges on the idea of political society as a solidaristic community, in which social cooperation is directed toward the

¹⁰⁵ Rawls, *Theory* pp. 253-4.

¹⁰⁶ Rawls, *Justice as Fairness* pp. 123-4; p. 126.

collective advancement of its members. We may, however, question whether this holds at the intergenerational level. After all, by virtue of the fact that generations are spread out across time, the prospects for reciprocal interaction of the kind envisaged by Rawls's argument for the difference principle are severely limited. Put another way, it is difficult to see how having distant descendants who will be considerably better off than us as a result of historical contingencies turns us into a means to an end. On the contrary, this might simply be an unavoidable feature of human societies. We may, of course, suggest that a feeling of solidarity might hold between proximate generations, but that this sense diminishes over time. This, of course, is highly problematic, because it would count against the selection of the difference principle within the general assembly interpretation of the original position. Simply put, if the types of reciprocity that hold between generations are limited, then the obligations which flow from such interaction will be equally restricted. This, then, brings us back to the circumstances of justice – and Rawls's contention that they represent the conditions which make social cooperation 'possible and necessary.' Rawls is therefore using the present time of entry interpretation to privilege a certain kind of cooperation within his theory: namely, reciprocal interaction. By way of contrast, those who favour the general assembly interpretation may be drawn toward a very different conception of social cooperation – or perhaps even, theories of justice which presuppose no such interaction. The debate between the standard contract model and its universalist analogue should, therefore, be seen as a moral rather than a methodological debate: the question is not how best to incorporate future generations into our deliberations, but what weight should be attributed to their interests. As we will now see, the proponents of a universal contract are in fact arguing in favour of full distributive equality across generations.

3.3 Equality and Intergenerational Justice

Contractualist theory is founded on a commitment to moral equality: it is for this reason that any initial agreement is directed toward specifying principles that could be justified to those who will be

subject to them. More particularly, Rawls's utilisation of the veil of ignorance serves to further this commitment, by neutralising any disparities in knowledge or bargaining power. It may, however, be the case that in the intergenerational context, the standard contract approach fails to live up to its egalitarian ambitions. Consider, for instance, Brian Barry's critical remarks in his 'Justice between Generations':

'For the whole idea – and the intellectual fascination – of 'justice as fairness' is that it takes self-interested agents, and, by the alchemy of the 'original position', forces them to choose principles of universal scope. In relation to subsequent generations, the postulate of self-interest is relaxed to allow concern for our successors, but this naturally limited sympathy is not forced by the logic of the 'original position' to be extended any further than it extends naturally. Our limited sympathies towards our successors are fed into the sausage-machine of 'justice as fairness' and returned to as duly certified obligations. We come seeking moral guidance and simply get our prejudices underwritten...'¹⁰⁷

The accusation, then, is that the present time of entry interpretation militates against our reaching genuinely impartial conclusions in our deliberations on intergenerational justice. Building upon this line of criticism, Stephen M. Gardiner has identified what he terms the 'close to the bone' objection to the standard contract approach. In outline, Gardiner argues that an inability to address questions of intergenerational justice undermines the appeal of the contract method itself. In part, this reflects the clear and obvious asymmetries in power which characterise the intergenerational context. More acutely, though, Gardiner suggests that the problems identified in the intergenerational realm are broadly analogous to those that obtain at the intragenerational level. Taken together, these observations lead to a troubling conclusion:

¹⁰⁷ Brian Barry, 'Justice between Generations' in *Liberty and Justice: Essays in Political Theory 2* (Oxford: Clarendon Press, 1991) p. 253.

'[These] considerations suggest two things. First, the intergenerational setting might raise issues very close to those that motivate many to be contract theorists in the first place. This suggests that a deep inability to overcome problems in these areas may undermine the appeal of the whole contract approach. Second, the situation arising between generations closely resembles those of central concern to contract theorists *even though the element of cooperation (as understood on the exchange model) appears to be absent*. Hence, the analogy introduces some doubt about whether the standard account of cooperation must or should play a central role in how we understand such situations.'¹⁰⁸

The implication, then, is that by focusing on cooperation, the standard contract approach cannot address the problems of the intergenerational context in a manner consistent with the fundamental values that underpin its appeal. In response to this, both Gardiner and Barry argue that an agreement with universal scope would be the obvious way of extending the contract approach to the intergenerational level in a manner that is consistent with a commitment to equality.¹⁰⁹ The central idea motivating universalism, then, is a moral one: future generations are comprised of people who warrant the same level of respect as our contemporaries, and we therefore have to seek their agreement within the original position. Notice that the idea of cooperation plays no role in this articulation of the contract approach. Instead, our obligations to future generations derive from a principle of equal concern for the rights and interests of people across time.

A principle of equal concern, however, is merely an abstract ideal which can, in practice, manifest in any number of forms. Indeed, were we to agree with Ronald Dworkin's analysis, then we could conceivably argue that nearly all theories of justice appeal to the idea of equal concern.¹¹⁰ We therefore need to say more about how and why the standard contract approach violates this norm. Put

¹⁰⁸ Stephen M. Gardiner, 'A Contract on Future Generations?' in in Axel Gosseries and Lukas H. Meyer (eds.) *Intergenerational Justice* (New York, NY.: Oxford University Press, 2009) p. 96.

¹⁰⁹ *Ibid.* pp. 114-5; Barry, 'Justice between Generations', pp. 253-4.

¹¹⁰ Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA.: Harvard University Press, 1978) pp. 180-3.

another way, what would be wrong – from the perspective of equality – in attending to the rights and interests of future people through the prism of a series of humanitarian obligations? In this respect, Barry's analysis is particularly instructive:

'The alternative route out of Rawls's difficulties is to pursue the logic of his own analysis more rigorously. This entails scrapping the part of the construction specifying that all people in the 'original position' are contemporaries and know that they are. We should now have to imagine that there is a meeting to decide on intergenerational relationships at which all generations are represented. Clearly the 'veil of ignorance' would be required to conceal from them which generation each of them belonged to. Otherwise, an early generation would always have the whip-hand over a later one in the negotiations.'¹¹¹

Here, then, we have a clear expression of why the present time of entry interpretation can be said to undermine the foundations of the contract method: making people aware of their generation of birth is comparable to informing them of their initial social position or native endowments. The present time of entry interpretation is therefore understood to violate a principle of *impartiality* within the original position. Recall that Rawls places his party behind a veil of ignorance to build impartiality into the structure of his initial agreement. The idea, in outline, is that mutually disinterested parties who are made ignorant of morally arbitrary facts about themselves and their society will select impartial principles of justice in order to further and promote their interests. Barry, though, draws attention to one respect in which Rawls does not follow through on his stated aim: in making the parties to the original position aware of a morally arbitrary fact about themselves, in the form of their being of the same generation. We might, then, be able to summarise the central argument motivating the universal contractualist position in a simple syllogism:

¹¹¹ Barry, 'Justice between Generations' pp. 253-4.

(1) When selecting principles of justice, we should exclude all morally arbitrary facts about ourselves and the societies we inhabit from our deliberations.

(2) Generation of birth is one such morally arbitrary factor.

(3) Therefore, when selecting principles of justice, we should exclude all knowledge of our generation of birth from our deliberations.

Notice, that Rawls does go some way to depriving his parties of facts about their generation of birth – they are unaware of the specific generation that they will be born into – but this does not translate into a general prohibition on knowledge relating to generational membership within the original position.¹¹² It is also worth stating that even if only one generation were to participate in the agreement, the very fact of making them unaware of their being contemporaries would ensure that the contract had intergenerational scope. Removing the present time of entry interpretation therefore has profound consequences for the contract approach and would appear to make good on a prior commitment to deliberative impartiality.

We may, however, begin to unpick this position by questioning whether or not the argument provided actually requires an intergenerational contract. On first impressions, the universalist argument seems eminently agreeable: after all, the key normative premise – that we ought to exclude morally arbitrary facts from our deliberations – is common to most forms of egalitarianism, and one's generation of birth is very clearly a morally arbitrary fact about oneself. On closer inspection, though, a problem can be discerned: this argument says nothing about the distributive scope of justice. To elaborate, were we to maintain that the obligations of distributive justice held only between contemporaries, then we could agree to the universalist argument, but continue to maintain that the obligations of

¹¹² Rawls, *Theory* p. 118.

justice were intragenerational in nature. On such an account, failing to apply full distributive equality to people of other *contemporary* generations would be a source of injustice, on the grounds that generation of birth is morally arbitrary. That, however, is as far as our egalitarian commitments would take us. Those proposing a universal contract, then, need to make an additional argument for extending the scope of justice to apply across generations. In this regard, two very different strategies can be identified. Firstly, it may be argued that the obligations of justice are owed to people as such, and extending the scope of our obligations across time is a natural corollary of this commitment. Following its use in the global justice literature, I will refer to this position as *non-relational egalitarianism*.¹¹³ At the same time, though, it may be the case that the obligations of justice are triggered only in specific contexts, but that – contrary to Rawls – this is not mutually advantageous cooperation, but membership of a given ethical community. Although these theories need not be communitarian in the narrow sense of the term, I will refer to them under the heading of *relational communitarianism*. As we will see, the latter of these two lines of argument represents a promising alternative to the reciprocal logic that informs Rawls's interpretation of the contract method.

Before proceeding any further, though, I would like to sound a word of caution. In the following section, we will employ concepts and arguments from the global justice literature in order to elucidate positions in the intergenerational context. It is, then, important to bear in mind that the focus of this enquiry is on intergenerational justice at the domestic or institutional level. Consequently, even though certain of the positions that I have in mind would conceive of the demands of justice in global, as well as transgenerational terms, for our present purposes we are only interested in their application to the intergenerational context.

i) Non-relational approaches

¹¹³ Andrea Sangiovanni, 'Global Justice, Reciprocity, and the State,' *Philosophy and Public Affairs* 35 (2007): 3-39 pp. 5-6.

Within contractualist thought, obligations of justice are triggered by specific forms of social interaction. Non-relational approaches dispute this, conceiving of justice as being owed to people irrespective of their institutional standing. Equality is therefore understood to be a foundational value with universal scope. Within the philosophical literature, this position receives its clearest articulation in the following remark by Charles R. Beitz on the subject of global justice:

'If the original position is to represent individuals as equal moral persons for the purpose of choosing principles of institutional or background justice, then the criterion of membership is possession of the two essential powers of moral personality – a capacity for an effective sense of justice and a capacity to form, revise, and pursue a conception of the good. Since human beings possess these essential powers regardless of whether, at present, they belong to a common cooperative scheme, the argument for construing the original position globally need not depend on any claim about the existence or intensity of international social cooperation.'¹¹⁴

Justice is therefore understood to have intergenerational scope by virtue of the intrinsic features of human beings. On first impressions, this is highly agreeable: after all, if we can conceive of two individuals with the same needs and interests, living at different times, then it would be strange to suggest that we should differentiate their claims on the basis of their institutional standing with respect to us. On the contrary, a natural way to approach any disputes that might arise between them would be to afford them parity in our deliberations. This, broadly speaking, is the position adopted by Barry in response to the problem of intergenerational justice.¹¹⁵ In summary, then, individuals – rather than institutions – are, on the non-relational account, assumed to be the appropriate subjects of any moral

¹¹⁴ Charles R. Beitz, 'Cosmopolitan Ideals and National Sentiment,' *Journal of Philosophy* 80(10) (1983): 591-600 p. 595.

¹¹⁵ Brian Barry, 'Justice Between Generations', in *Liberty and Justice: Essays in Political Theory 2* (Oxford: Clarendon Press, 1991), 242-58. See also, Reglitz, 'Practice-Independence of Intergenerational Justice'.

enquiry, and our moral and political obligations should reflect this. Institutions, then, are seen as the instruments through which justice can be pursued, rather than the site of our justice-based obligations.

Although this conception has a certain moral and intellectual appeal, it is far from obvious how a concern for the needs and interests of future generations – or, for that matter, the global poor – translates into a demand for distributive equality. To elaborate, by pursuing equality we are attempting to bring about a certain relationship between past, present, and future: even if the basic needs of posterity are met, we have not fulfilled our obligations until their needs and interests are met to the same extent as ours. We might, therefore, want to enquire as to what the relevant features of human individuals are that should lead us to pursue distributive equality between them, rather than, say, sufficiency. In this way, one can reasonably dispute the central contention of non-relational egalitarianism that equality is a fundamental value within our deliberations. More particularly, the non-relational approach can lead to problematic conclusions. Consider, for instance, the – highly likely – possibility that future generations will enjoy significantly higher living standards than we do, by virtue of technological advancement. Although this would result in significant disparities between individuals, it is far from clear how this constitutes an injustice; on the contrary, it is simply the inevitable consequence of a historical process. Moreover, even if we accept that it is a source of injustice, it is difficult to see how it could be addressed without levelling down: we might, for instance, be required to stymie the process of technological change, or to deliberately squander resources. Although very few non-relational egalitarians would advocate such methods, it does hint at a wider problem for the theory: namely, that it builds motivations which approximate to envy into our deliberations. So much to say that non-relational egalitarianism, at least as it is presently constituted, rests on a controversial conception of equality which many would reject.

It may, of course, be the case that the universal contract could take a non-egalitarian form. Indeed, if my analysis is correct, then it is eminently probable that the principles selected in such circumstances

would engender obligations of a far less demanding nature than full distributive equality. Moreover, we can retain the intrinsic qualities of human beings as our starting point but reach very different conclusions if we appeal to the idea that people have basic needs which ought to be met as a matter of justice. What I have in mind here, then, is a universal contract which gives rise to a form of intergenerational sufficientarianism. This would avoid many of the counterintuitive conclusions of non-relational egalitarianism and accord with certain of our intuitions concerning future generations. Unfortunately, it is difficult to see how a principle of sufficiency would ever be selected within the confines of the contract approach. Put simply, many of the ideas which undergird contractualism are at variance with the idea of sufficiency as a distributive principle. The most obvious of these being why rational parties looking to maximise their individual shares of primary goods would be minded to pursue sufficiency when they could be much better off. Indeed, this issue is particularly acute when examined within the intergenerational context for the simple reason that the disparities in wealth are so vast: some generations will have to subsist on the breadline whilst others will command vast amounts of wealth. Reaching a principle of sufficiency within the original position would, therefore, only be possible if we were to radically alter the motives and fundamental nature of the parties to it. Indeed, for all of its failings, non-relational egalitarianism was, at the very least, broadly consistent with the nature of the parties to any initial agreement. Given the nature of the original position, it is difficult to see how non-egalitarian principles of justice could be selected. Furthermore, a principle of intergenerational sufficiency could also lead to deeply counterintuitive conclusions. Consider, for instance, an example in which an earlier generation intentionally squander resources, leaving their successors on the threshold – thereby meeting their distributive obligations whilst needlessly restricting prospects of their descendants. The central problem, however, for any non-egalitarian form of non-relational contractualism consists in the fact that the original position is structured in such a way as to lead to broadly egalitarian conclusions. Unless we are willing to amend some of the fundamental elements of the contract approach, we would therefore be advised to look elsewhere for a universalising strategy.

ii) *Relational communitarianism*

With the non-relational approach dealt with, we now turn to a relational argument for extending the scope of justice to our distant successors: namely, the broadly communitarian idea that they are owed justice by virtue of their being members of the same community as us. Although this idea is presented as being a communitarian one, its appeal is arguably far wider. One can, for instance, see aspects of this approach in Rawls's characterisation of political society as a 'fair system of social cooperation over time, from one generation to the next.'¹¹⁶ More specifically, an argument of this nature forms the basis of David Miller's liberal nationalist thesis, according to which our obligations to posterity arise from the historical sacrifices of our forebears: their having 'toiled and spilt their blood' in service to the nation obliges us to continue the national project.¹¹⁷ For now, however, we will focus on a more explicitly communitarian account, in the form of Janna Thompson's discussion of lifetime-transcending interests. In summary, lifetime-transcending interests are a specific type of interest which relate to both our posthumous concerns and the historical projects of our forebears. Consequently, these lifetime-transcending interests can generate a rationale for intergenerational cooperation which does not depend on mutual advantage: instead, we cooperate with future generations to promote the various projects that we care about. More than simply providing the dead with a vehicle for the continuation of their projects, however, it is argued that such lifetime-transcending interests provide future generations with a fertile intellectual and cultural inheritance in which to develop and act upon their own projects. This, Thompson suggests, motivates an 'objective interest' in the 'continued existence and survival...of the heritage [we have] obtained from the past,' leading us to identify our own interests with those of political society – thereby motivating certain intergenerational distributive obligations.¹¹⁸ We can therefore conceive of our having reasons to

¹¹⁶ Rawls, *Political Liberalism* p. 14.

¹¹⁷ David Miller, *On Nationalism* (Oxford: Clarendon Press, 1995) p. 23.

¹¹⁸ Janna Thompson, 'Identity and Obligation in a Transgenerational Polity' in Gosseries and Meyer, *Intergenerational Justice* pp. 37-8; 43-5.

extend the scope of justice into the future, most obviously with respect to promoting these objective interests in the survival of our community and the collective projects that characterise it.

Of course, the appeal to community raises two difficulties from the intergenerational perspective: firstly, our sense of community tends to be temporally limited. We might, for instance, conceive of our community as extending over a few generations, rather than across time itself. Secondly, it may very well be the case that future generations ascribe to a different set of moral and political beliefs than we do. We might, therefore, have very little reason to extend the scope of our distributive obligations to distant generations on account of their being members of our ethical community. In this regard, Thompson's position is arguably more resilient. As she notes:

'We can reasonably assume that people of the more remote future will have lifetime-transcending interests and will make lifetime-transcending demands. Given the importance of these interests to human life, as we understand it, it is morally necessary to make this assumption. Even remote generations can be regarded as successors, as participants in a never-ending chain of relationships in which each generation fulfils moral requirements in respect to its predecessors and successors. The people of the remote future may have different moral ideas and different political institutions, but so long as they are in the above sense our successors, we have a duty not act in ways that might undermine the conditions that need to exist if they are to maintain their institutions and pursue their interests.'¹¹⁹

The idea, then, correlates broadly with Edmund Burke's conception of political society as being a compact between 'those who are living, those who are dead, and those who are [yet] to be born.'¹²⁰ In brief, our community is, by its very definition, a transgenerational entity and to the extent that we

¹¹⁹ *Ibid.* p. 47.

¹²⁰ Edmund Burke, *Reflections on the Revolution in France* (London: Penguin, 1968) p. 195.

identify with it, we have reasons to extend our distributive obligations into the future. In this way, there is a potential mechanism which may allow for a genuinely intergenerational contract. Future generations are therefore members of our political community and even our distant descendants are, in a sense, our successors. Consequently, if we are possessed of a desire to promote and sustain a just society over time, granting posterity parity in our deliberations may be seen as a requirement of justice.

3.4 Community, Reciprocity, and Posterity

We may, then, have reason to extend the scope of our obligations on account of a commitment to furthering the communal projects which underpin political society. On such a reading, the demand for intergenerational equity arises from a need to specify fair terms of collaboration. Notice, however, that unlike the standard contract approach, this cooperation does not have to be mutual; the only prerequisite is our participation in the 'never-ending chain of relationships in which each generation fulfils moral requirements in respect to its predecessors and successors.' In this final section, I intend to contrast this approach with a reciprocity-based view of our obligations to future generations, suggesting that the latter is preferable to the former. In contrast with the community-based view, the reciprocal approach holds that our obligations to posterity are based on historical or expected contribution to a process of intergenerational savings. To summarise, then, the reciprocal approach locates our justice-based obligations to posterity in our having received certain goods from our predecessors and the expectation that we will make an equitable and equivalent transfer to our successors. Obligations of intergenerational justice are therefore directed toward the development, realisation, or sustenance of certain key public goods: just institutions, environmental integrity, and the like. Within the context of the community-based approach, however, obligations of justice attempt to cultivate equitable relations between past, present, and future in order to further certain communal projects – in particular, the values and aspirations of political society itself. In what follows, I will

argue that the reciprocal approach is better positioned to capture our basic intuitions about intergenerational justice.

To that end, it may be worth elaborating on the obvious contrast that can be drawn between the two approaches. On the one hand, the community-oriented account locates the obligations of justice within a moral framework in which we are required to preserve certain goods and resources to allow for the continued existence of our community. Our obligations, then, turn upon the values and mores which characterise an ethical community. On the other, the reciprocal approach locates our obligations in the production and preservation of a specific set of goods and resources. Crucially, such preservation takes place for its own sake. In other words, on the reciprocal account we do not have to conserve natural resources because they are conducive to the survival of our community; on the contrary, we might be obliged to preserve them because of their intrinsic importance within a cooperative endeavour over time. This distinction is significant, because the community-based approach is arguably less resonant with our intuitions than the reciprocal one. Were we to discuss the problems of resource depletion over time, for example, its effect on certain values would likely be a tangential concern, as compared to the harms visited on future generations and the inherent unfairness of squandering a valuable resource in this way. Suppose, however, that the depletion of a given resource had no bearing on the life prospects of future generations or the survival of our community. In that case, might there still be something objectionable in our using up vast amounts of resources? Instinctively, I would argue that there was. By acting wastefully and using more than our fair share of the resource in question, we would be depriving future generations of various opportunities to utilise it in the pursuit of their own projects, even if their material prospects were unaffected. In this respect, the reciprocal approach is better placed to explain certain of our intuitions within the intergenerational context. That, of course, is not to say that the preservation of community is wholly irrelevant from the perspective of intergenerational justice, but the concerns it addresses are of a largely subsidiary nature in the context of issues such as climate change.

A further concern facing the community-based view relates back to the original position: how, in other words, could we translate its fundamental insights into the initial agreement. After all, the parties in the original position are looking to maximise their shares of a relatively specific set of primary goods. There is, then, an obvious corollary between the reciprocal view – according to which the principles of justice exist to regulate the institutions and practices which generate and safeguard those goods – and the contract method as it is presently constituted. There is, of course, the possibility that the values which the community-based view seeks to promote relate back to primary goods in some way: we might posit that a political culture which embodies certain values is a necessary condition for the protection and promotion of basic liberties, for instance. That, however, draws the communal approach toward its reciprocal analogue: certain values, on this understanding, are worthy of being promoted for instrumental reasons – namely, their connection to primary goods. Indeed, one can see how certain cultural practices could, on this account, be rebranded within the language of reciprocity as a form of contribution. The problem with pursuing a unificatory strategy of this nature is that as both approaches become increasingly enmeshed, the scope for a universal contract is greatly reduced. After all, the core features of the intergenerational context – the temporally distant nature of the parties, asymmetric power relations, and so on – militate against a universal contract founded on reciprocity. The communal approach can, therefore, be rendered intelligible within the context of contractualist theory. Unfortunately, this would come at the expense of the universalising ambitions which initially motivated the theory. For that reason, I am inclined to believe that a universal contract of the kind described would only be possible if we were willing to alter some of the fundamental aspects of contractualist theory. For that reason, I now propose to outline the major obstacles to a theory of intergenerational justice grounded in reciprocity.

3.5 Three Problems of Intergenerational Reciprocity

Given my endorsement of an approach founded on reciprocity, it would be fitting to conclude this chapter with a very brief summary of the problems encountered when attempting to apply this idea to the intergenerational context. To that end, I want to identify three recurring problems within the wider literature which need to be met if an account of our obligations to posterity is to be understood in reciprocal terms: scope, mutuality; and justification. Before beginning, though, it should be noted that due to the asymmetrical relations between past, present, and future generations, reciprocal interaction will necessarily take an indirect form. Intergenerational cooperation, then, involves one generation benefitting another who, in requital, confer broadly equivalent benefits on a third generation. More will be said about this in Chapters 4 and 5, but for now we simply need to be aware of the problems this may engender from the perspective of justice.

i. The Problem of Scope

Since indirect reciprocal interaction occurs between proximate generations, we may question the extent to which it can meet the hard problem of intergenerational justice. Joseph Heath, for instance, has developed an extensive theory of intergenerational cooperation founded on intergenerational overlap.¹²¹ That, however, won't satisfy our stated aim of arriving at a theory which extends between non-overlapping generations. Furthermore, even if we can conceive of our obligations as being owed to our contiguous, non-overlapping successors, there is still the problem of the intergenerational time bomb. The problem of scope can, therefore, be understood to comprise two distinct questions: firstly, can a theory of intergenerational justice founded on indirect reciprocity generate obligations toward

¹²¹ See, Joseph Heath, 'Intergenerational Cooperation and Distributive Justice', *Canadian Journal of Philosophy* 27(3) (1997): 361-76; Heath, 'The Structure of Intergenerational Cooperation'.

non-overlapping generations; and secondly, can it specify sufficiently robust regulations to safeguard our more distant successors?¹²²

ii. The Mutuality Problem

As we have seen, the contract approach is predicated on the idea that principles of justice regulate mutually advantageous social cooperation. That, however, raises the question of how reciprocal interaction satisfying these conditions can obtain in the radically asymmetrical intergenerational context. Indeed, the benefits of intergenerational cooperation are understood to flow in only one direction: from earlier generations to their successors. That, then, would suggest that conditions of mutuality do not exist at the intergenerational level; which, in turn, would considerably limit the scope and content of our obligations to future generations. If, therefore, we want to use indirect reciprocal interaction as a grounding for our obligations toward posterity, then we need to find some way of responding to the mutuality problem.¹²³

iii. The Justificatory Problem

Finally, there is the problem of justifying obligations that are generated through indirect reciprocal interaction. To elaborate, consider Barry's celebrated toffee apple example: in it we are given a toffee

¹²² See, e.g., Gardiner, 'A Contract on Future Generations' pp. 83-5; Anja Karnein, 'Climate Change and Justice Between Nonoverlapping Generations', *Global Justice: Theory Practice Rhetoric* 8(2) (2015): 43-65 p. 52.

¹²³ See, e.g., Attas, 'A Transgenerational Difference Principle' pp. 189-92; Gustaf Arrhenius, 'Mutual Advantage Contractarianism and Future Generations', *Theoria* 65(1) (1999): 25-35 pp. 33-4; Klaus Mathis, 'Future Generations in John Rawls' Theory of Justice', *Archives of Law and Social Philosophy* 95(1) (2009): 49-61.

apple by a benevolent stranger. How is it, Barry asks, that accepting and enjoying this gift generates an obligation to buy a toffee apple for a different stranger?¹²⁴ There is, then, a question as to how indirect reciprocal interaction could engender obligations in the first instance – and that raises questions as to the justifiability of any savings principle within the confines of the contractualist approach. Within the literature, this is frequently cited as *the* canonical refutation of indirect reciprocity as the basis of our intergenerational obligations.¹²⁵

In responding to these challenges, I intend to prioritise the mutuality problem. The belief being, that the lack of mutuality between generations contributes to the problems of scope and justification. If, then, we can advance a cogent line of response to the mutuality problem, then the problems of scope and justification will become less daunting. Consequently, we will begin Chapter 4 by attending to the mutuality problem, before going on to deal with the problems of scope and justification in the latter half of the chapter

3.6 Conclusion

In conclusion, this chapter has set out to establish four things. Firstly, it has been argued that the disputes between competing interpretations of the original position reflect moral, rather than methodological concerns. Simply put, the present time of entry interpretation serves to build the idea of mutually advantageous cooperation into the structure of Rawls's agreement. Secondly, it has been suggested that a universal agreement is most readily intelligible within the context of a wider commitment to community – and to preserving the projects and values which characterise political

¹²⁴ Brian Barry, 'Justice as Reciprocity' in *Liberty and Justice: Essays in Political Theory 2* (Oxford: Clarendon Press, 1991) p. 232.

¹²⁵ See, e.g. H.P. Visser 't Hooft, *Justice to Future Generations and the Environment* (Boston, MA.: Kluwer Academic Publishers, 1999) pp. 61-2; Richard Vernon, *Justice Back and Forth: Duties to the Past and Future* (Toronto: Toronto University Press, 2016) pp. 198-9.

society. Finally, it has been suggested that the reciprocal approach which is implied in the standard formulation of the contract method is preferable to the community-based logic of the universal account.

In addition, three problems affecting the reciprocal approach have been identified, and it has been suggested that the mutuality problem is the most significant obstacle to a contractualist theory of intergenerational justice. I will therefore respond to it in the next chapter with reference to my theory of intergenerational mutuality. With the mutuality problem resolved, I will then attend to the problems of scope and justification. In this way, advancing a theory of intergenerational mutuality will be shown to be the key to solving the three problems of intergenerational reciprocity.

4. The Intergenerational Social Contract

The intergenerational realm is characterised by asymmetrical relations between past, present, and future generations. Simply put, earlier generational cohorts are able to affect the prospects of their successors without fear of requital; by way of contrast, future generations are unable to exert an equivalent influence on their distant predecessors. In turn, this militates against the kind of like-for-like cooperation which, within the confines of the modern social contract approach, engenders the demands of distributive justice. The absence of mutuality between generations would therefore imply that our intergenerational obligations are of a more general, humanitarian nature – and owed to people as such, rather than having a specific recipient. Such a conclusion would, however, be at variance with the notion of a just savings principle – which, following John Rawls's discussion of justice between generations in *A Theory of Justice*, has emerged as the most practicable means of incorporating the interests of future generations into our distributive deliberations in the present. According to such an approach, the realisation and maintenance of just institutions over time enjoys an absolute priority over measures designed to make the contemporary worst off as well off as possible. In other words, it is only once the interests of posterity have been attended to that distributive justice can be realised at the intragenerational level.¹²⁶ A principle of just savings, in other words, has the appearance of a justice-based obligation to future generations. In order to substantiate the problem facing the extension of a contractualist theory of justice to the intergenerational realm, consider the following antinomy comprising three plausible intuitions:

¹²⁶ John Rawls, *Theory* pp. 250-8.

(1) Firstly, that those obligations associated with distributive justice arise within the context of mutually advantageous cooperation.

(2) Secondly, that principles of distributive justice are required to constrain intragenerational expenditure, with a view to securing the continuation of a just institutional framework.

(3) And finally, that the interaction between discrete cohorts at the intergenerational level cannot be characterised by mutuality.

To express this another way: principles of justice are needed to regulate the intergenerational context, but they presuppose some form of mutually advantageous social cooperation which cannot obtain between generations. To resolve this problem, one of these three contentions therefore has to be rejected.

Arguably the most radical line of response would involve the repudiation of the claim that discussions of distributive justice presuppose some form of institutional interaction. By expanding the scope of our distributive obligations, such an approach effectively sidesteps the problems associated with the asymmetry between past and future generations. As we have argued, though, such a move has a transformative effect on the contract method itself – recasting it in a manner devoid of those reciprocal features that go some way to explaining its appeal. By way of contrast, a more modest strategy is implied by the rejection of the second intuition. On such an account, our obligations to posterity are of an altogether more limited nature than those which obtain between contemporaries. Indeed, if this is the case then, as David Heyd has noted, the just savings principle can be reasonably likened to the duty of assistance between peoples at the global level: both reflect a broad commitment to justice as a value worthy of

promotion, whilst stopping well short of being obligations *of* justice.¹²⁷ The principal concern with such an approach revolves around whether or not humanitarian duties of the kind envisaged can grant sufficient weight to the interests of our descendants in our distributive deliberations. Could, in other words, a set of general obligations owed to people as such constrain the realisation of justice at the intragenerational level in the way that the just savings principle attempts to do. If not, then we are left with the third and final intuition – that relations between non-overlapping generations can never be characterised by mutuality. Although the rejection of this belief may, at first, seem counterintuitive, this chapter will argue that such a move represents a promising line of response to the question at hand.

To that end, I will appeal to a concept of intergenerational mutuality in order to illustrate the reciprocal nature of cooperation at the intergenerational level. The idea, in outline, is that our posthumous interests are more important to us than we generally acknowledge. Consequently, there is a sense in which we are vulnerable to future generations and need to cooperate with them in order to realise certain essential collective benefits. On such an account, the just savings principle serves to regulate this cooperation in an equitable manner. In order to develop this case, we will begin by expanding on the concept of reciprocal interaction and its relationship to mutuality in Section 1. In Section 2, I will examine the absence of mutuality from the intergenerational context, before expounding my own theory of intergenerational mutuality in Section 3. In Section 4, I will identify a second source of mutuality within the contract account, and in Sections 5 and 6 I will respond to the problems of scope and justification. What will ultimately emerge from this exercise is a compelling line of response to the mutuality problem

¹²⁷ David Heyd, 'A Value or an Obligation? Rawls on Justice to Future Generations' in in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (Oxford: Oxford University Press, 2009) pp. 183-5.

which serves to outline of an approach to questions of intergenerational justice which will be subject to greater scrutiny in the following chapter.

4.1 Reciprocity

In order to bring clarity to our discussion of the mutuality problem it would be helpful to begin by saying something about reciprocity as it informs the contractualist analysis. After all, it is not immediately obvious why mutuality and reciprocity should have to go hand in hand. To illustrate this by way of an example from recent history, the nuclear ambitions of the United States and the Soviet Union gave rise to a relationship characterised by mutuality without any concomitant reciprocal content. Likewise, we can conceive of transactions in a market economy as being reciprocal, but they are not necessarily mutual. We therefore need to say something about how the two concepts converge within the contractualist analysis. Consider, for instance, the concept of mutually advantageous social cooperation: it is distinguished from mere social cooperation by virtue of the fact that the benefits will be felt by all parties. Similarly, we could consider the mutual obligations that might obtain between citizens, or the mutual burdens that parties may have to shoulder in the pursuit of a societal goal. What is ultimately being described in these examples is an idea of correspondence: the benefits or burdens being visited upon one party are correlative with those experienced by others. This idea also informs our understanding of reciprocity, albeit in a subtly different way. To return to the example of a market transaction, it is reciprocal to the extent that the value of the good being purchased corresponds to its price. Similarly, at the societal level, we might understand a principle of reciprocity to be a one which rewards people in a manner that is commensurate with their contribution. Within the contract method, however, reciprocity is used to inform the initial agreement: the parties to the original position are understood to be equal participants in

a cooperative venture who are owed a fair return on their contribution. On the one hand, then, the parties are understood to have made corresponding contributions to the scheme, whilst on the other, they are owed a return which is commensurate with their input. In this way, the logic of both reciprocity and mutuality becomes entwined within the structure of the original position. It is for this reason that the apparent absence of mutuality from the intergenerational context is troubling from the perspective of a reciprocity-based approach: the parties cannot be assured that other generations are making a correlative contribution to the proposed scheme. Despite being distinct concepts in their own right, mutuality and reciprocity can therefore be understood to converge within the contract method.

With the relationship between reciprocity and mutuality now clarified, we can move on to a broader explication of reciprocity. After all, reciprocal interaction can take multiple forms, and we need to be clear about the focus of our present enquiry. On a particularly narrow reading, for instance, a principle of reciprocity might be understood to give rise to a demand for remuneration in direct proportion to one's marginal contribution. If, however, reciprocity is understood in broader terms, incorporating social norms, then the resultant obligations will be radically different. Before going any further into our enquiry, then, it will be necessary to distinguish different forms of reciprocity from one another and to identify the form of reciprocity that will provide the focal point for this chapter.

For our present purposes, two salient distinctions can be drawn between different forms of reciprocity. Firstly, between *direct* and *indirect*— and whether or not reciprocal interaction is understood in exclusively like-for-like terms or can extend to third-party transactions. In a direct arrangement, reciprocity involves making a like-for-like return on an initial contribution to the group or individual responsible. By way of contrast, an indirect approach will see the

return on the initial benefaction being made to a third party. To illustrate, a recognisable instance of direct reciprocity occurs whenever an individual buys a friend or colleague a drink in the expectation that, at some point in the future, the other party will return the favour in kind. A comparable indirect reciprocal arrangement takes the form of the first party buying a drink for the second party on the understanding that the recipient will go on to buy a drink for the third party. In this way, indirect reciprocal interaction can be likened to the pattern of bequests and inheritance which characterises the intergenerational realm.

The second distinction to be drawn is between *strong* and *weak* forms of reciprocity. Reciprocal interaction is generally understood in terms of deriving some return from an initial sacrifice or making good on an earlier benefaction. The contrast between strong and weak forms of reciprocity hinges upon how the return is to be appraised. On a weak understanding, reciprocal interaction is understood in self-interested terms, and motivated by the expectation of future benefits. Strong reciprocal schemes, by way of comparison, are characterised by a willingness to reward fair behaviour and to penalise actions which violate behavioural norms, even when doing so will require resources to be expended in a manner that is incommensurate with any present or future benefits.¹²⁸ To expand upon this, imagine a scenario in which we receive good service in a restaurant to which we will never return. On a weak reciprocal analysis, we would have no reason to tip, since we would be incapable of extracting any future benefits from the transaction. There would, however, be a strong reciprocal reason to tip, since it would represent our making a fair return to the waiting staff. Within the context of strong reciprocity, therefore, the expectation of future benefits is of limited importance as compared to the role that this form of interaction plays in strengthening and upholding social norms.

¹²⁸ Herbert Gintis, Samuel Bowles, Robert Boyd, and Ernst Fehr, 'Moral Sentiments and Material Interests: Origins, Evidence, and Consequences' in Herbert Gintis, Samuel Bowles, Robert Boyd, and Ernst Fehr (eds.) *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life* (Cambridge, MA.: The MIT Press, 2005) pp. 3-13.

Within the context of our present enquiry, our focus will be on strong, indirect reciprocity. More specifically, we will be concerned with whether or not the form of indirect collaboration which obtains between generations can engender justice-based obligations. The issue at hand will, therefore, be whether or not the present generation's inheritance of a stock of capital from their predecessors requires them to make an equivalent transfer to their descendants. By limiting our enquiry to strong reciprocal norms, we are also able to overcome many of the issues which ordinarily characterise the intergenerational realm. For instance, the problem of a first generation who have no ancestors from whom to receive any inheritance from need not arise. On the contrary, the earliest generation can be understood to initiate these practices because they have an interest in bringing about fair intergenerational norms. Similarly, as and when generational cohorts differ in size the content of any distributive obligations can change to reflect this. Put another way, if one generation has far more descendants than their predecessors, then they would be required to leave more than they inherited – so as to respect the underlying distributive principle. In this way, we can concentrate squarely on the issue of whether a distributive imposition of the kind envisaged could be justified through reference to indirect reciprocity.

4.2 The Problem of Intergenerational Justice

Within the confines of a contractualist approach, principles of justice serve to regulate the terms of mutually advantageous social cooperation. The idea, in outline, is that the parties to any agreement need to collaborate to further their interests. At the same time, though, these individuals will make competing claims on the goods obtained in this way. Accordingly, a unique set of distributive norms are required to apportion the benefits of this cooperative

endeavour in an equitable manner.¹²⁹ The obligations of justice, in other words, are triggered by a particular kind of reciprocal interaction. It is in this regard that the intergenerational realm is problematic from a contractualist perspective. Simply put, the types of interaction which obtain between generations are not conducive to mutuality. To elaborate, earlier generations are able to further the interests of their descendants, but subsequent generations cannot materially benefit their dead predecessors. The benefits of any cooperation between discrete generations are therefore understood to move unidirectionally. If this is an accurate portrayal of the intergenerational context, then earlier generations would have little – or perhaps no – reason to save; their participation in any savings process would yield few benefits and engender appreciable burdens. We therefore find ourselves in a position where mutually advantageous cooperation looks to be impossible. If that is the case, though, then a principle of intergenerational saving could not be selected in the original position. The lack of mutuality at the intergenerational level therefore militates against there being any obligations of justice between generations.¹³⁰

Faced with this problem, three possible lines of response can be identified. Firstly, we could attempt to excise the idea of mutually advantageous social cooperation from the contractualist framework. Such a move would have a transformative effect on the contract approach, but may prove necessary if the more conventional strategies can be found to be wanting. Secondly, we may acknowledge that we have no justice-based obligations to future people but suggest that a set of humanitarian obligations can regulate the intergenerational domain. In this way, any principles of intergenerational saving can be likened to those duties of assistance which obtain

¹²⁹ Rawls, *Theory* pp. 3-10.

¹³⁰ See, e.g., Stephen M. Gardiner, *A Perfect Moral Storm: The Ethical Tragedy of Climate Change* (New York, NY.: Oxford University Press, 2011) pp. 160-9.

between peoples at the international level.¹³¹ And finally, we may argue that intergenerational cooperation is, in actual fact, characterised by mutuality. By appealing to the idea of strong, indirect reciprocity, such a response will try to demonstrate that earlier generations are benefitted by collaborating with their descendants. For our present purposes, we can set aside the first, universalising strategy. Instead, we will focus on the second and third lines of response. By offering a brief outline of the shortcomings of a humanitarian approach, I hope to place further emphasis on the desirability of a reciprocity-based account. Then, by demonstrating the potential for mutually advantageous cooperation over time, this chapter will attempt to dissolve the problem of intergenerational justice.

In light of the problems inherent in attempting to extend the contractualist method to the intergenerational context, one may be tempted to follow Heyd in recasting our obligations to future generations in general, humanitarian terms. On such an account, our responsibilities to posterity can be likened to those duties of assistance which obtain between peoples at the international level. A process of intergenerational savings would therefore aim to transfer sufficient levels of capital to our successors to allow for the preservation of just institutions. Our intergenerational obligations, then, are understood to be principles *securing* justice rather than principles *of* justice.¹³² Were that the case, though, then we would expect the resultant obligations to be somewhat weaker than those envisaged by the just savings principle. It is, for instance, difficult to see how a set of humanitarian obligations could enjoy priority over the difference principle in the way that the just savings principle does. More acutely, though, there are notable differences between the intergenerational and international contexts, not least with respect to the level of influence earlier generations are able to exert over later ones. Indeed,

¹³¹ Heyd, 'A Value or an Obligation' pp. 183-4; John Rawls, *The Law of Peoples* (Cambridge, MA.: Harvard University Press, 1999) p. 107.

¹³² *Ibid.* pp. 183-7.

returning to Stephen Gardiner's analysis, the inability to extend our distributive obligations to the intergenerational realm may, in actual fact, undermine the appeal of the contract method itself. After all, much of the intuitive appeal of contractualism derives from its ability to overcome the asymmetries in power which characterise agreements in other social contract approaches. An inability to address the unavoidable asymmetry which characterises the intergenerational realm – and which can have a profound and pervasive impact on the life prospects of later generations – is, therefore, especially problematic.¹³³ Moreover, as Heyd concedes, generations tend to have a solidaristic attachment to their descendants and their predecessors.¹³⁴ Even if this intuition were to be illusory, the sense of cooperation between proximate generational cohorts demands further investigation. It may, for instance, be the case that there are intermediary norms that fit somewhere between the full demands of justice between contemporaries and our general, humanitarian obligations. Indeed, I think this represents a more plausible interpretation of the just savings principle: generations may not be obliged to extend the full demands of distributive equality to the intergenerational context, but they are required to restrain their intragenerational pursuit of it to ensure that future generations receive an equitable bequest. In the rest of this chapter, it will be argued that this feeling of solidarity between generations has a concrete basis. By appealing to the idea of strong, indirect reciprocity we can overcome the difficulties identified in this section by demonstrating that intergenerational cooperation *is* mutually advantageous. Earlier generations would therefore be willing to comply with a savings process because they would derive benefits from it. My strategy, then, is to block the mutuality problem at source. To put this another way, it is my contention that the parties to the original position would select a principle of intergenerational

¹³³ Stephen M. Gardiner, 'A Contract on Future Generations' in *Intergenerational Justice* eds. Gosseries and Meyer p. 96.

¹³⁴ Heyd, 'A Value or an Obligation' p. 184.

savings for mutually disinterested reasons alone. The obstacles to the just savings principle being selected will therefore not arise.

4.3 Indirect Reciprocity, Mutuality, and the Social Bases of Self-Respect

To that end, we will begin by looking at indirect reciprocity in greater detail. As has been mentioned, Rawls's justification of the just savings principle hinges upon our accepting that cooperation of some form obtains at the intergenerational level. In this regard, the Rawlsian intuition can be understood to proceed from the conception of political society as a fair system of cooperation over time.¹³⁵ With respect to this claim, it is possible to identify a pattern of inheritance and bequests which characterise the intergenerational context in cooperative terms. With the exception of the first and final generations, each cohort finds itself the recipient of a historical transfer and a benefactor with regard to their immediate descendants. A theory of indirect reciprocity attempts to formalise this relationship and use it as the basis for our intergenerational distributive obligations. In this way, the realisation and maintenance of just institutions over time can be interpreted as a transgenerational project to which each generation is expected to contribute. The idea would therefore be that the present generation's willingness to participate in a savings process is a function of their having benefitted from it. That, however, raises the question as to how the first generation could benefit from this process. After all, they will receive no benefits from their predecessors but will still be required to save for their successors. Faced with such a prospect, the parties to the original position would – assuming that maximin is the correct strategy – take this perspective into account when specifying an intergenerational savings rate. Of course, such an analysis presupposes that earlier generations can derive no benefits whatsoever from intergenerational cooperation. That, however, is to

¹³⁵ Rawls, *Theory* p. xv.

overlook the posthumous interests and ongoing projects that characterise the intergenerational domain. To elaborate, the parties to any agreement are aware of their having a conception of the good and it is reasonable to suggest that, in the overwhelming majority of cases, this will entail our having what Janna Thompson refers to as 'lifetime-transcending interests.'¹³⁶ By this, it is meant that our finding value in certain objects and features of the world can – and very often does – provide us with interests which extend beyond our respective lifespans. Suppose, for instance, that someone was to devote their life to curating a collection of valuable art and artefacts: in those circumstances, it is highly likely that the person would care about the survival – and possible expansion – of that collection long after their death. After all, it is likely that one of the reasons which possessed this individual to begin her collection was a desire to bring these intrinsically valuable items together for people to appreciate in perpetuity. In outline, then, a great many people have mutually disinterested reasons to care about the terms of interaction between generations – and even those who don't may be so disposed behind a veil of ignorance. The idea that we have lifetime-transcending interests which future generations can either promote or set back therefore introduces the possibility that mutuality may obtain at the intergenerational level.

In seeking to expand on this claim, Samuel Scheffler's argument in *Death and the Afterlife* represents a useful point of reference. In order to demonstrate the importance of posthumous interests within our overall conceptions of the good, Scheffler proposes two related thought experiments exploring the impact of impending extinction on the things we value. In the first, simply entitled the 'doomsday' scenario, we are informed that thirty days after our death an asteroid will collide with the Earth, destroying the planet and all life on it. Being made aware

¹³⁶ Janna Thompson, 'Identity and Obligation in a Transgenerational Polity', in *Intergenerational Justice* eds. Gosseries and Meyer pp. 33-4.

of this, Scheffler argues, would horrify most people. From this, it is inferred that much of what we value in life matters to us for reasons other than our being able to experience them.¹³⁷ More significantly, though, it is argued that the doomsday scenario would precipitate an existential crisis within our lives: the projects and activities which had previously given us some kind of purpose would become either wholly meaningless or severely compromised.¹³⁸ The second thought experiment, borrowed from P.D. James's novel *The Children of Men*, envisages a future in which humanity faces extinction through mass infertility, ensuring that human civilisation will die out within a century. This, Scheffler suggests, would trigger a similar response to the doomsday scenario, leading to 'widespread apathy, anomie, and despair'.¹³⁹ This is significant for two reasons: firstly, unlike the doomsday scenario, it would not involve anyone we loved or cared about dying prematurely; secondly, though, it would entail the untimely demise of innumerable human projects. This has profound implications for a theory of justice, because it suggests that our conceptions of the good are, in some way, dependent upon there being a future extending well beyond our lives. It should, of course, be noted that less doom-laden responses to these thought experiments have been posited: Susan Wolf, for instance, has suggested that, after a period of anhedonic despair we would eventually return to the task of attempting to live a fulfilling life.¹⁴⁰ What is relevant to our present discussion, though, is the thought that an imminent extinction event could prompt a wholesale re-evaluation of our individual conceptions of the good. It implies that, far from having no reasons to care about the future, mutually disinterested individuals could be minded to consider the fate of posterity in the original position.

¹³⁷ Samuel Scheffler, *Death and the Afterlife* ed. Niko Kolodny (New York, NY.: Oxford University Press, 2013) pp. 18-20.

¹³⁸ *Ibid.* pp. 25-7.

¹³⁹ *Ibid.* pp. 37-40.

¹⁴⁰ Susan Wolf, 'The Significance of Doomsday', in *Death and the Afterlife* pp. 127-8.

In attempting to flesh out this line of argument, I would want to suggest that similar concerns obtain in ordinary, non-apocalyptic circumstances. Consider, for instance, our curator – who is concerned, among other things, about the survival of her collection of art and artefacts into the distant future. Now suppose, for the sake of the argument, that she were to be made aware of the fact that shortly after her death, a particularly malevolent individual would divide her collection up and unceremoniously sell it off for monetary gain. Such knowledge, I would suggest, would have a profoundly deleterious effect on our curator – possibly to the extent of destabilising her entire conception of the good. This, in large part, is down to the fact that the wrongdoing occurs after the curator has died: had it happened in her lifetime, then at least she could intervene to either stop or mitigate the process. What I am suggesting, then, is that the asymmetry in power relations between past, present, and future generations has a flip side which consists in the vulnerability of our projects and posthumous interests to acts of non-compliance by future generations. The curator, for example, has dedicated her life to building a collection on the understanding that it would survive well into the future: her overarching project in life is, then, dependent upon future generations either maintaining or contributing to the collection that she has amassed. Scheffler notes as much in his 2018 work *Why Worry About Future Generations*, in which relations between the present generation and its successors are taken to have reciprocal content: '[t]he term “reciprocity”’, writes Scheffler, ‘is appropriate in this context because our relation to future generations is one of genuine mutual dependence, even if the form taken by our dependence on them is different from the form taken by their dependence on us. Each side is dependent on, and so is vulnerable to, what the other does or what happens to the other.’¹⁴¹ There is, in other words, a form of mutuality between generations. Moreover, I believe that we can analyse this mutual dependence in terms of a collective action problem – of the sort envisaged within the contractualist framework. To put

¹⁴¹ Samuel Scheffler, *Why Worry About Future Generations?* (Oxford: Oxford University Press, 2018) p.75.

this somewhat crudely, our successors want us to save on their behalf; and we want them to preserve – or, at the very least, show respect toward – our projects and posthumous interests. In order, then, to achieve the outcomes that both want, the present generation and its successors need to strike an intergenerational bargain. Now, within the context of the original position, the idea of an agreement between generations is out of the question, due to the present time of entry interpretation. It is, however, possible for members of the present generation to specify a principle to regulate this practice by asking themselves which principle they would want previous generations to have abided by.¹⁴² The thinking, then, is that by selecting a savings principle that we would have wanted our predecessors to have followed, we might secure the compliance of future generations in upholding our projects and respecting our posthumous interests. In this way, it may be possible to envisage the just savings principle being selected by mutually disinterested parties, without any amendments to the structure of the original position.

There is, of course, an obvious rejoinder to the argument I have just outlined. Namely, how could these diffuse and varied interests be incorporated into the contract approach? Recall that the parties to the original position are looking to maximise their share of a relatively specific list of primary goods: they are not, in other words, simply trying to promote whatever interests we may wish to attribute to them. Indeed, it is for this reason that Thompson suggests that lifetime-transcending interests are simply too particular to be admissible within the original position.¹⁴³ Elsewhere, Eric Brandstedt has suggested amending the list of primary goods to include ‘sustainability of values’.¹⁴⁴ My argument, however, would be that if our lifetime-

¹⁴² Rawls, *JF* p. 160.

¹⁴³ Janna Thompson, *Intergenerational Justice: Rights and Responsibilities in an Intergenerational Polity* (New York, NY: Routledge, 2009) p. 52.

¹⁴⁴ Eric Brandstedt, ‘The savings problem in the original position: assessing and revising a model’, *Canadian Journal of Philosophy* 47(2) (2017): 269-289.

transcending interests are as animating as I have suggested, the contract method is already capable of attending to these concerns. Indeed, I would want to appeal to the social bases of self-respect as a potential candidate for bringing these considerations to bear within the confines of the orthodox contract approach. To elaborate, the social bases of self-respect are, by Rawls's reckoning 'perhaps the most important primary good' and are defined as having two aspects: firstly, a person's sense of their own value – including the idea that their conception of the good is worth carrying out; and secondly, having confidence in their ability to fulfil their intentions.¹⁴⁵ If the Schefflerian analysis is correct, then the idea that our posthumous interests might be disregarded would, to my estimation, threaten the first feature of the social bases of self-respect: it would have the potential to undermine our conception of the good – or, at the very least, to raise severe doubts as to its worthiness. This, as Rawls suggests, would have particular significance to the parties in the original position:

When we feel that our plans are of little value, we cannot pursue them with pleasure or take delight in their execution. Nor plagued by failure and self-doubt can we continue in our endeavors. It is clear then why self-respect is a primary good. Without it nothing may seem worth doing, or if some things have value for us, we lack the will to strive for them. All desire and activity becomes empty and vain, and we sink into apathy and cynicism. Therefore the parties in the original position would wish to avoid at almost any cost the social conditions that undermine self-respect.¹⁴⁶

Notice that we don't have to concur with Scheffler's conclusions to see how our relations with future generations might affect the social bases of self-respect. Indeed, all we need to acknowledge is that our conceptions of the good will, very frequently, give rise to posthumous

¹⁴⁵ Rawls, *Theory* p. 386.

¹⁴⁶ Loc. cit.

interests which place us in a relationship of mutual dependence with future generations. In this way, it is possible to move from the idea of mutually disinterested parties specifying principles of justice at the intragenerational level to a concern for the relations between past, present, and future generations. Put another way, if earlier generations are minded to cooperate with their successors, then the mutuality problem does not obtain within the intergenerational context.

If the mutuality does obtain at the intergenerational level, then we can return to the original position and ask whether or not the just savings principle would be selected. Recall that the just savings principle specifies that sufficient levels of capital are to be transferred to the next generation to allow for the realisation and maintenance of just institutions, with a further stipulation that no generation should leave less to their forebears than they received through the savings process.¹⁴⁷ We therefore need to consider what features of the original position would lead to this being selected over a more straightforward principle of sufficiency. As we noted in Section 1, when mutuality obtains we can be assured that others are making an equivalent contribution to the scheme in question: in the intergenerational context this means that past, present, and future generations are all fully participating parties within the scheme of intergenerational cooperation and are owed an equitable return. To demonstrate why this would lead to the selection of the just savings principle rather than a principle of sufficiency, consider a generation who only transfers enough capital to their successors to maintain just institutions – thereby leaving their descendants worse off than they were. This would clearly violate those reciprocal norms which underpin the contract method: by allowing the earlier generation to derive disproportionate benefits relative to the burdens that the scheme entails, a principle of sufficiency would be rejected on the grounds that it is not impartial. In this context, Rawls's

¹⁴⁷ See, e.g., John Rawls, *Justice as Fairness: A Restatement* ed. Erin Kelly (Cambridge, MA: Harvard University Press, 2001) pp. 159-60.

discussion of the strains of commitment and the psychological harms involved in having to shoulder disproportionate burdens relative to others is especially relevant: inequitable terms of social cooperation, Rawls suggests, can have a pernicious effect on an individual's prospects to the extent of undermining their capacity for self-respect.¹⁴⁸ In this regard, there is an obvious symmetry between the just savings principle on the one hand, and the practice of honouring posthumous interests on the other: both serve to promote and safeguard a specific primary good in the form of the social bases of self-respect. Consequently, the parties to the original position would select the just savings principle on mutually disinterested grounds. To put this another way, allowing individual generational cohorts to leave less to their successors than they received would be seen as an unreasonable gamble within the original position: by assuming the position of the least advantaged, the parties would gravitate toward a principle which maximised their individual share of primary goods – in this case, the just savings principle. We can, therefore, understand the just savings principle as being reflective of the idea of reciprocity within the intergenerational context and as a further guarantor of the social bases of self-respect.

4.4 Mutuality and Stability

Before proceeding to look at the problem of scope, I would like to draw attention to a second source of mutuality within the contractualist approach. In outline, it is my contention that an appeal can be made to the stability of a conception of justice – the requirement that institutions organised in accordance with the principles of justice can generate their own support – to introduce a secondary form of mutuality into the theory.¹⁴⁹ I refer to this as 'secondary' because

¹⁴⁸ *Ibid.* pp. 127-8.

¹⁴⁹ *Ibid.* pp. 230-1.

within the original position, a test for stability follows the selection of principles: if they are found to be wanting in this regard, then they are rejected.¹⁵⁰ Such a commitment may engender obligations of the sort envisaged in two closely related ways: firstly, in instances where making intergenerational transfers inculcates and reinforces a ‘sense of justice’ among contemporaries; and secondly, when making a bequest to our successors serves to stabilise a conception of justice in the long-run.¹⁵¹ There may, in other words, be a sense in which the inability of future generations to make a like for like return is immaterial from the perspective of justice. If, for example, intergenerational saving strengthens certain norms at the intragenerational level, then the practice would be mutually advantageous. Compliance with a savings process could, then, be to the benefit of the present generation, meaning that there is an additional form of intergenerational mutuality which can be incorporated into the contract method.

To expand upon this, it is worth noting that stability is a necessary condition for the enjoyment and appreciation of those primary goods whose distribution is the principal concern of those parties behind the veil of ignorance. Indeed, as Rawls himself notes, a stable conception of justice is a one which is ‘perspicuous to our reason, congruent with our good, and rooted...in the affirmation of the self’ – or, to put it another way, one in which the chosen principles of justice create a social environment in which we are free to pursue our goals in conditions characterised by mutual respect between citizens.¹⁵² Given the future-oriented concerns engendered by a commitment to stability, it thereby follows that intergenerational transfers need to be informed by such considerations. Suppose, for instance, that prior generations had saved at a particularly low rate, so as to maximise their intragenerational wellbeing. Not only would such a state of affairs militate against the cultivation of a sense of justice in the

¹⁵⁰ *Ibid.* pp. 125-6.

¹⁵¹ *Ibid.* pp. 230-1; pp. 397-405.

¹⁵² Rawls, *Theory* p. 436.

contemporary generation, but it would also have a deleterious effect on the bonds of mutual respect which are essential to a well-ordered political society. In this respect, there may be a parallel to be drawn between Rawls's later argument for the difference principle as a guarantor of the least advantaged group's self-respect, and this argument for the just savings principle; both are couched in explicit commitments to reciprocity, solidarity, and an overarching concern that inequities in the distribution of goods may have pernicious consequences with respect to the ability of individuals to develop and pursue their own, individual conception of the good.¹⁵³ An approach to intergenerational justice which emphasises considerations relating to the stability of a conception of justice therefore enjoys considerable overlap with an account of our intergenerational obligations grounded in indirect reciprocity. By incorporating strong reciprocal interaction into the theory, however, we are able to respond to the charge that this form of interaction cannot be mutual. On the contrary, instances of non-compliance are to the detriment of present and future generations. Consequently, the decision to adopt a principle of just saving would be a rational course of action for those mutually disinterested parties in the original position.

4.5 Indirect Reciprocity and the Problem of Scope

As discussed in the previous chapter, the problem of scope is of particular relevance to this enquiry: we are, after all, focusing explicitly on the hard problem of intergenerational justice and looking at what is owed to non-contemporaneous future generations. We may, therefore, be able to discern a clear and obvious challenge to theories of intergenerational justice grounded in indirect reciprocity. As has been noted, the focus of such a theory is on securing

¹⁵³ John Rawls, *Justice as Fairness: A Restatement* ed. Erin Kelly (Cambridge, MA: Harvard University Press, 2001) pp. 122-6.

equitable terms of cooperation between proximate generational cohorts. Thus, although we can envisage duties and obligations to our distant descendants, the most onerous demands of intergenerational justice begin and end with our immediate successors. This gives rise to the problem of so-called 'time bomb' policies – wherein the deleterious effects are deferred for a number of generations, so as to ensure that our nearest descendants are unaffected. Although such hypotheticals may sound far-fetched, many of the most salient issues within the intergenerational realm – such as climate change – could, quite accurately, be characterised as time bombs of the sort envisaged. The problem, as Gardiner observes, is that contract theories which are grounded in mutual advantage appear to lack the conceptual resources to condemn acts of the kind envisaged.¹⁵⁴ Imagine, for instance, that the present generation wilfully diminished the prospects of all people born 150 years in the future to their immediate benefit. Since they would still be able to discharge their justice-based obligations to their successors, the argument goes, there would be no injustice on such an account. All things being equal, then, we may have reasons for scepticism with respect to the scope of the obligations envisaged.

In replying to this, it is worth stressing that, at the bare minimum, earlier generations will have humanitarian obligations to their distant descendants. Consequently, a policy which intentionally restricted the prospects of future generations below a certain threshold would be worthy of condemnation from the perspective of justice. For the sake of the argument, then, let us assume that the extent of the imposition is not so grave as to breach any humanitarian obligations, but enough to depart from standards of mutual respect between persons. In that case, a great deal would depend upon the temporal proximity of the generations – and the extent to which they are mutually dependent on one another. As I have demonstrated, cooperation between proximate generations would be regulated by something along the lines of a just

¹⁵⁴ Gardiner, 'A Contract on Future Generations' pp. 103-5.

savings principle. This, on my account, gives rise to the most demanding justice-based obligations which obtain in the intergenerational context. Intergenerational mutuality and indirect reciprocal cooperation are not, however, limited to relations between contiguous generational cohorts. Indeed, on the account I have put forward it would be wholly plausible to suggest that our justice-based obligations could extend over many centuries. We could, therefore, conceive of our obligations as moving gradually from the maximal standard of the just savings principle toward the humanitarian minimum. That, however, strikes me as an intuitively plausible way to think of the demands of intergenerational justice: our extremely distant successors *are* as removed from our present circumstances as members of remote tribes in this day and age. A theory of justice grounded in indirect reciprocity therefore has a number of resources to draw upon when faced with problems that take the form of an intergenerational time bomb. More specifically, though, the idea of intergenerational mutuality resolves the hard problem by necessarily extending the obligations of justice to non-contemporaneous generations. In this way, resolving the mutuality problem has overcome the problem of scope.

4.6 Intergenerational Mutuality and the Problem of Justification

Having attended to the problems of mutuality and scope, all that remains is to offer a brief account of how intergenerational mutuality resolves the problem of justification. Recall that Brian Barry attempted to draw attention to the apparent absurdity of indirect reciprocity by asking us to imagine ourselves receiving a toffee apple from a benevolent stranger. How, asks Barry, could our acceptance and enjoyment of the toffee apple engender a correlative duty to buy toffee apples for strangers? To develop this point, Barry envisages a set of circumstances in which receipt of a toffee apple would generate distributive obligations: if, for instance, there was an established practice of handing out toffee apples to strangers then, having taken and

enjoyed one, it would be appropriate to claim that the recipient ought to make an equivalent contribution – as in the case of people buying drinks in rounds. That, however, would still fail to account for why our obligations were owed to a third party.¹⁵⁵ In responding to the justificatory challenge, we therefore need to meet two objections to indirect reciprocity: firstly, we need to demonstrate that these transfers are part of a wider distributive scheme; and secondly, we need to know why our obligations are owed to a third party.

By virtue of the fact that we have demonstrated the existence of an extensive scheme of intergenerational reciprocity, the first part of the justificatory challenge can be dismissed outright. Instead, I want to focus on the second question raised in this context: why, in other words, are our obligations owed to third parties in the intergenerational context? To explain this, I would like to introduce Axel Gosseries's distinction between *ascending* and *descending* modes of indirect reciprocity. To elaborate, imagine three contiguous generations: A, B, and C. Suppose now, that A – the earliest generation in this sequence – transfers certain benefits to their immediate successors, B. On the descending model, this transfer would place Generation B under an obligation to make an equivalent transfer to their successors, C. By way of contrast, the ascending model operates in the opposite direction: Generation B may promote the interests of Generation A, and this is understood to oblige Generation C to act similarly toward their predecessors.¹⁵⁶ So, within the scheme we have described, transferring capital goods to the next generation would accord with the descending model. Conversely, the ascending model is manifest in the practice of honouring the posthumous interests of our forebears. The two models can therefore be understood to combine in the wider system of intergenerational reciprocity. Consequently, by advancing a theory of intergenerational mutuality, we are able

¹⁵⁵ Barry, 'Justice as Reciprocity' pp. 232-3.

¹⁵⁶ Axel Gosseries, 'Three Models of Intergenerational Reciprocity' in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (New York, NY.: Oxford University Press, 2009)

to identify an indirectly reciprocal scheme of cooperation between generations which resolves the justificatory dilemma.

4.7 Conclusion

In conclusion, it has been argued that obligations of justice, grounded in the idea of mutually advantageous cooperation, can obtain at the intergenerational level. Accordingly, the contractualist method can be brought to bear on the question of justice between generations. As we have seen, the perceived impossibility of mutuality at the intergenerational level presents the most significant obstacle to such an account. By expounding a concept of intergenerational mutuality, however, I have demonstrated that participation in an indirect reciprocal scheme yields benefits at the intragenerational level. More specifically, I have suggested that the types of cooperation which obtain between generations could be analysed as a bargain in the original position. To put this another way, the mutually disinterested parties in the original position would willingly select a principle of intergenerational saving. Complying with the associated norms of intergenerational saving is, therefore, in the interest of the present generation. Consequently, Rawls's contention that the parties to the original position can assume historical compliance is unproblematic. Indeed, with these amendments made to the contractualist account, the characterisation of the just savings principle as 'an understanding between generations to carry their fair share of the burden of realizing and preserving a just society' is an appropriate one. Moreover, the resultant theory has demonstrated that it can resolve the concomitant problems of justification and scope. Intergenerational cooperation, then, may take an indirect form, but it nevertheless gives rise to a unique set of distributive obligations between discrete generational cohorts. In this way, an approach to the question of justice between

generations grounded in indirect reciprocity is capable of offering a coherent and compelling response to the problem of intergenerational justice.

5. Intergenerational Mutuality: Objections and Responses

In the previous chapter, a theory of intergenerational mutuality was presented as a solution to those problems which militate against the extension of the contract method to the intergenerational context. To summarise, by envisaging the relationship between discrete generations as a one characterised by mutuality, we can address the three problems identified at the end of Chapter 3 – those of justification, scope, and mutuality. The concept of intergenerational mutuality therefore underpins the thesis as a whole. It would, then, be appropriate to give some consideration to the objections which might be raised against the theory. In this concluding chapter, we will focus on two sources of scepticism concerning intergenerational mutuality: doubts about posthumous interests and concerns surrounding historical non-compliance and enforcement. In order to flesh these sceptical intuitions out into something more concrete, we will begin by setting out four arguments against intergenerational mutuality in Section 1, before responding to each in Sections 2, 3, and 4, respectively. The first of these will focus on the appropriateness or otherwise of posthumous interests from the perspective of justice. They may, for instance, be unreasonable, in which case granting them any standing in the original position would be inherently problematic. The second line of argument I want to address concerns the commensurability of posthumous interests with those goods generated by the wider process of intergenerational savings. The point of contention in this instance is whether the idea of an intergenerational bargain is plausible when the goods moving in either direction are so qualitatively different. In responding to both of these claims, particular emphasis will be placed on the relationship between posthumous interests and the social bases of self-respect, thereby allowing for further exposition of this aspect of the theory. The third and fourth points of contention that will be examined in this chapter are closely

related and concern the resilience of the proposed scheme in the face of non-compliance. To elaborate, we have, thus far, proceeded on the assumption that generations will comply with their obligations to one another. In practice, however, generations may fail to fulfil their obligations and we need to consider how the indirect reciprocal scheme could recover from acts of historical non-compliance for two reasons: firstly, widespread non-compliance has the potential to undermine the scheme in its entirety; and secondly, the only enforcement mechanism that is available to later generations – disregarding the projects and posthumous interests of their non-compliant predecessors – would precipitate the collapse of the reciprocal arrangement. In responding to both of these issues in Section 4, I will look to expand upon the idea that each generation has an interest in bringing about fair intergenerational norms. That, of course, raises the separate question of whether or not the proposed scheme is actually characterised by reciprocity. After all, if generations have grounds to bequest goods to their successors irrespective of what they received from their predecessors, then this calls the reciprocal credentials of the theory into question – and in Section 5.5, we will attend to this issue. Responding to these objections will therefore serve two purposes: firstly, it will allow us to tie up any conceptual loose ends within the thesis, and secondly it will bring additional layers of nuance to the theory. What will hopefully emerge, then, is a more thorough and detailed understanding of mutuality in the intergenerational context.

5.1 Four Objections

To provide some insight into the four objections that form the principal subject of this chapter, it would be expedient to offer a brief sketch of the central argument in favour of intergenerational mutuality as outlined in the broader thesis:

(1) Many people will have posthumous interests which, in some cases, will be inextricably bound to their conception of the good.

(2) This means that within any political society, there will be a population whose goals and projects are dependent on the compliance of future generations. To this extent, relations between past, present, and future generations are characterised by mutuality: just as earlier generations can promote the interests of their successors by participating in a process of intergenerational savings, later ones can benefit their forebears by furthering their posthumous interests.

(3) For that reason, those norms which obtain at the intergenerational level can be enforced. The parties to the original position will therefore be minded to select a principle of intergenerational savings.

There are, then, three key moves in the argument: firstly, the recognition of posthumous interests themselves; secondly, the implied symmetry between the posthumous interests of earlier generations and the benefits conferred via intergenerational saving; and thirdly, the idea that this allows us to punish historical non-compliance and provides a mechanism by which the principles selected in the original position can be enforced. From this, we can identify four major points of contention which demand further discussion:

i) Firstly, do posthumous interests warrant the kind of recognition within the original position that we have afforded them? Here, what I have in mind is something along the lines of the idea – prominent in the writings of Thomas Paine and Thomas Jefferson – that it is objectionable

for earlier generations to impose their projects on their successors.¹⁵⁷ To elaborate, the theory of intergenerational mutuality set out in this thesis appeals to the idea of posthumous interests to motivate the parties in the original position to consider the prospects of future generations. The notion, in outline, is that the parties to the agreement will see an intergenerational savings process as a means of furthering their posthumous interests. That, however, leaves open the question of whether or not these motives are appropriate within the original position. After all, if we were being uncharitable, this could be characterised as allowing earlier generations to bind future generations into their projects via the mechanism of an intergenerational savings process. Instead, it may be better to think of posthumous interests as a feature of human psychology which, like envy or spite, is unreasonable from the perspective of justice and so plays no role in the selection of principles.¹⁵⁸ The first charge, then, is that a theory of intergenerational mutuality depends upon unreasonable motivations being made admissible within the original position: if this is the case, then the argument in favour of intergenerational mutuality would be fatally compromised.

ii) Secondly, are the goods flowing in either direction within the proposed scheme commensurate? To put this another way, those capital goods being transferred to our successors via the just savings principle are qualitatively different to any benefits they can confer on us by furthering our posthumous interests. We therefore need to consider how these goods relate to one another and whether there can be any correspondence between them. It may, for instance, be that we need to give thought to how one set of goods might be exchanged with the other.

¹⁵⁷ See, e.g., Thomas Jefferson, 'Letter to James Madison, Paris, 6th September, 1789' in *Thomas Jefferson: Writings* ed. Merrill D. Peterson (New York, NY: Literary Classics, 1984); Thomas Paine, 'The Rights of Man' in *Rights of Man, Common Sense, and Other Political Writings* ed. Mark Phillips (New York, NY: Oxford University Press, 1995) pp. 91-2.

¹⁵⁸ Rawls, *Theory* pp. 123-4.

Obviously, if the goods are so dissimilar that intergenerational trades of this nature cannot take place, then the proposed reciprocal scheme would prove to be illusory.

iii) Thirdly, we need to further consider how non-compliance would affect the scheme: might it be possible to salvage something from our forebears, would the scheme wholly collapse, or could it be restarted? This is particularly relevant in cases where prior generations have failed to fulfil their obligations in a widespread or deleterious manner – to the extent that their legacy represents a threat to indirect reciprocal interaction between generations.

iv) Fourthly, we need to examine the enforcement mechanisms available within the intergenerational context. The most obvious of these would see the posthumous interests of non-compliant generations being disregarded. However, as with the previous issue, this runs the risk of either undermining or bringing about the collapse of the indirect reciprocal scheme.

The theory of intergenerational mutuality is, then, vulnerable to multiple objections. In responding to these, we will begin by looking at the role of posthumous interests in the original position in Section 2, before moving on to consider the commensurability of the goods generated by indirect intergenerational reciprocity in Section 3. Owing to the overlap between the two issues, the questions of non-compliance and enforcement will be attended to simultaneously in Section 4. This process will ultimately draw attention to two features of the theory of intergenerational mutuality: firstly, the centrality of the social bases of self-respect within it, and secondly, the scheme's ability to reconstitute itself in conditions of widespread non-compliance.

5.2 Posthumous Interests and the Original Position

In defending the concept of posthumous interests within the original position, it would be helpful to begin by appealing to an idea present in Rawls's own work – as well as the wider liberal canon – namely, that generations ought to enjoy a significant degree of autonomy with respect to one another. Indeed, to quote directly from *A Theory of Justice*: '[A]ll generations have their appropriate aims. They are not subordinate to one another any more than individuals are and no generation has stronger claims than any other.'¹⁵⁹ There are, then, limits on what discrete generations can demand of one another and the just savings principle – at least as Rawls envisages it – exists to specify this. By way of contrast, the just savings principle as envisaged by the theory of intergenerational mutuality can also be thought of as a mechanism through which earlier generations can promote their posthumous interests. To reiterate, the earlier generation bequests a stock of capital goods to their successors and the later generation reciprocates by honouring the projects and posthumous interests of their predecessors. We therefore need to consider whether or not this modified savings principle has the potential to encroach upon the autonomy of future generations. After all, on an uncharitable reading, the savings process could be recast as a mechanism by which earlier generations can bind posterity into the projects by foisting benefits on them. This is of a particular significance, because Rawls is quick to exclude motivations which militate against fair terms of social cooperation – such as envy or spite – from his conception of rationality.¹⁶⁰ It may very well be, then, that posthumous interests are a common feature of our lived experience which, nevertheless, are rendered inadmissible by the structure of the original position. Obviously, if that were true, then the argument for intergenerational mutuality would fall at the first hurdle. We therefore

¹⁵⁹ Rawls, *Theory* p. 257.

¹⁶⁰ *Ibid.* pp. 123-30.

need to further examine posthumous interests, their role in the original position, and their relationship to the just savings principle.

To begin to respond to this line of criticism, it would be helpful to offer a brief summary of how posthumous interests are factored into the original position within the wider theory of intergenerational mutuality. Recall that within Rawls's formulation of the original position, the parties are looking to maximise their share of primary goods, and I argued that posthumous interests relate specifically to the social bases of self-respect. The idea, in outline, is that some people will have conceptions of the good which are inextricably linked to their posthumous interests. Consequently, the idea that the parties to the original position would be wholly indifferent to the circumstances of future generations can be called into question. Moreover, since the parties to the agreement are unaware of their specific conception of the good they will be mindful of the fact that they could find themselves in this position. In this way, the vulnerability of particular conceptions of the good to the actions of future generations would be a universal concern within the original position. The parties to the original position are not, therefore, looking to promote their posthumous interests *per se*. Instead, their concerns about the relationship between the present generation and its successors turns upon the need to secure the social bases of self-respect. With this in mind, the question facing the parties to the agreement is how to bring about a relationship between the present generation and its successors in which those conceptions of the good which entail significant posthumous interests are afforded the appropriate degree of respect. Posthumous interests are not, therefore, something to be promoted in their own right within the original position. Instead, they are afforded standing within our deliberations due to their being a component of the social bases of self-respect. It would, then, be more appropriate to characterise the resulting principle of

intergenerational savings as one specifying equitable terms of cooperation – such that past, present, and future generations can interact on the basis of mutual respect.

With the role of posthumous interests in the original position now established, we can turn our attention to the question of how burdensome this arrangement would be for posterity. The first – and arguably most salient – thing to say is that attempting to bind future generations into our projects would run contrary to the goal of the savings principle envisaged by the theory of intergenerational mutuality. After all, the intergenerational savings process is undertaken to ensure that the burdens of social cooperation over time are distributed equitably and in a manner that is consistent with mutual recognition and respect between generations: attempting to shackle posterity in the way envisaged would clearly violate this core intuition. Instead, the idea embedded within intergenerational mutuality sees the projects of earlier generations as being worthy of respect – to the extent that they should be afforded appropriate weight in our deliberations. Notice, however, that this does not imply that posthumous interests have to necessarily trump the rights or interests of future generations. To elaborate on this, consider the example of the curator from Chapter 4. Suppose, for the sake of the argument, that she has built up an extensive collection of jewellery, ornaments, and other assorted artefacts made from an especially rare metal. Now, for the sake of the argument, suppose that future generations were to consider melting her collection down for monetary gain. In these circumstances, we can clearly see that their actions violate the standards of mutual respect that characterise a theory of intergenerational mutuality. Conversely, were the rare metal an essential technological component that could save millions of lives, then the rationale behind melting the collection down would be altogether more justifiable – albeit in a way that acknowledges that however great the gain, something of value would be being lost. Obviously, a theory of justice is not designed to adjudicate on specific cases of this type, however, this example

hopefully captures the core intuition behind the idea. As before, the goal of the savings process is to safeguard the social bases of self-respect within the intergenerational context. By linking posthumous interests to the social bases of self-respect we can, therefore, incorporate them into the theory in a way which avoids affording asymmetrical power to earlier generations and prevents onerous burdens being placed on posterity.

5.3 Posthumous Interests, The Just Savings Principle, and the Social Bases of Self-Respect

The theory of intergenerational mutuality conceives of discrete generations as participants in a series of broadly symmetrical exchanges: earlier generations save on behalf of their successors, and later ones respect and, in some cases, actively promote the projects and interests of their forebears. It is, however, an unavoidable fact that these interactions involve very different goods flowing in either direction. We therefore need to consider how the stock of capital goods that the present generation will pass on to their immediate descendants relates to those posthumous interests that later generations can promote on behalf of their antecedents. The concern, in brief, is that the goods being exchanged between generations are simply too dissimilar to ground the distributive obligations in question. Indeed, the comparatively diffuse nature of ‘lifetime-transcending’ interests is what ultimately led Janna Thompson to abandon her attempt to incorporate them into the original position.¹⁶¹ My strategy in this regard has been to incorporate posthumous interests into the social bases of self-respect. That, however, only serves to prompt the question of *how* posthumous interests relate to the social bases of self-respect. The proposed correspondence between the just savings principle, posthumous interests, and the social bases of self-respect could, for instance, be radically dissimilar in either

¹⁶¹ Janna Thompson, *Intergenerational Justice: Rights and Responsibilities in an Intergenerational Polity* (New York, NY: Routledge, 2009) p. 52.

instance. The argument in favour of intergenerational mutuality would therefore lose much of its cogency. Consequently, in this section, I want to argue that the goods flowing in either direction are of a broadly comparable nature from the perspective of justice. To that end, I intend to establish a correspondence between the just savings principle and the honouring of posthumous interests as they relate to a person's capacity to develop and act upon a conception of the good. In this regard, my argument against the claim can be understood to take the insights found in the previous section and integrate them into the analysis of the just savings principle in the previous chapter.

Within the confines of the contract approach, primary goods are understood to have universal appeal to rational parties. At its most basic, the idea is that however divergent and incommensurable our individual conceptions of the good will be, we will require primary goods if we are to develop and act upon them.¹⁶² It is for this reason that the goods transferred through an intergenerational savings process and those that might be promoted by honouring the posthumous interests of our forebears must ultimately relate back to the metric of primary goods. After all, those goods could, ultimately, be put to manifold ends: what matters from the perspective of justice is that they possess the essential characteristic of being desirable irrespective of one's ultimate goals. In arguing for the just savings principle over its competitors, I suggested that an appeal to the social bases of self-respect could explain its being selected ahead of rival savings principle in the original position. Similarly, I have argued that the honouring of posthumous interests are best understood through the prism of the social bases of self-respect. That, however, is not sufficient evidence to state that there is an equivalence between the goods being exchanged in this way: one, for instance, might be intimately bound to the social bases of self-respect, whilst the other is only tangentially related to it.

¹⁶² Rawls, *TJ* p. 54.

Consequently, we need to recap the arguments linking both to the same primary good. Recall that posthumous interests are associated with the social bases of self-respect owing to their centrality to certain conceptions of the good: were the curator to assume that her collection would be disbanded at the point of her death, then it is highly likely that her major lifelong project would, in some sense, be destabilised. The honouring of posthumous interests therefore enables people to develop and act upon conceptions of the good that may otherwise appear futile. The just savings principle, by way of comparison, was linked to the social bases of self-respect on the grounds that generations ought to bequeath a comparable amount of capital to their successors as they received from their forebears. To elaborate on this, it was suggested that an earlier generation who acted in a particularly profligate manner – to the detriment of their successors – would visit certain psychological strains on their descendants whose ability to develop and act upon their own conceptions of the good could be compromised. There is, then, a congruence between the two arguments: honouring the posthumous interests of prior generations and transferring an appropriate stock of capital to later ones both relate to the social bases of self-respect via their impact on the capacity of individuals to develop and act upon a conception of the good. To that extent, these seemingly disparate goods may be comparable from the perspective of justice.

There is, however, is an obvious rejoinder to this line of argument: as I have acknowledged, the number of people whose conceptions of the good relate so closely to their posthumous interests could be fairly limited, whereas in the argument for the just savings principle, an entire population may be affected. I, however, would disagree with this analysis: if we are to take Rawls's commentary on the strains and commitment seriously, then we ought to be alive to the harms that can visited on individuals in this way; indeed, they may be especially acute if the

group in question is few in number.¹⁶³ Imagine, for instance, that the burdens of the previous generation's imprudence fall disproportionately on the worst-off group within the successor generation: if anything, the psychological harms being envisaged may be rendered more deleterious by virtue of their relative isolation from the rest of society. More pointedly, from the perspective of the original position, numerical impact ought to have no standing: however rare the features of a given conception of the good may be, the parties – being unaware of their overarching goals and projects in life – would have to err on the side of caution and assume that they may, in some way, be dependent on future generations who will respect their posthumous interests. For this reason, I contend that despite the qualitative differences between the goods in question, they are, from the perspective of justice, convergent. The honouring of posthumous interests and the transfer of capital goods in accordance with the just savings principle both serve to safeguard the social bases of self-respect, allowing individuals to develop and act upon conceptions of the good that would otherwise prove unattainable or futile. For this reason, I conclude that we are warranted in viewing the transfers between generations as symmetrical from the perspective of justice.

5.4 Historical Non-Compliance and Enforcement

Having dealt with posthumous interests, we can now proceed to examine those concerns relating to historical non-compliance and enforcement. In order to explicate the theory of intergenerational mutuality, we have, thus far, assumed that each generation will be fully compliant with the demands of justice. In practice, however, generations may fall short of their obligations, and this is significant in two ways. Firstly, we need to know what effect non-compliance would have on the scheme itself: it may, for instance, have the potential to

¹⁶³ See, e.g., Rawls, *JF* pp. 127-8.

undermine or perhaps even bring an end to reciprocal interaction between generations. Secondly, we need to consider what methods of enforcement later generations could employ to sanction historical non-compliance. This latter question is particularly troubling because the obvious method of enforcement available to subsequent generations involves their refusing to honour the posthumous interests of their non-compliant forebears; however, that has the potential to collapse the scheme in its entirety. After all, one of the central ideas within the theory of intergenerational mutuality is that generations are, at once, recipients of a bequest and the guarantors of posthumous interests: if the present generation cannot fulfil these roles, then the chain of obligation linking past, present, and future generations breaks down. This is of consequence from the perspective of the mutuality problem because, as Heyd notes, the parties to the original position require assurances that acts of non-compliance will be met with credible sanctions.¹⁶⁴ Consequently, the absence of viable enforcement mechanisms within the intergenerational context threatens the very idea of intergenerational mutuality. This is especially concerning because the sanctions available to later generations become increasingly drastic as non-compliance becomes more and more pervasive. It is for this reason that I have chosen to tackle these two issues simultaneously: in cases of widespread non-compliance, the only viable sanctions available to later generations threaten the reciprocal scheme itself. In order to respond to this, I will begin by looking at instances of limited non-compliance – where viable methods of enforcement are available – before moving on to examine wholesale non-compliance. In those instances, I will concede that the only sanctions available to later generations will bring about the collapse of the scheme. However, I will then argue that this is not a problem because those generations who have to employ such sanctions would restart the

¹⁶⁴ David Heyd, A Value or an Obligation? Rawls on Justice to Future Generations' in in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (Oxford: Oxford University Press, 2009) p. 180.

scheme. Consequently, credible mechanisms of enforcement are available to subsequent generations, and the theory of intergenerational mutuality can meet Heyd's challenge.

We will begin, then, by looking at how the scheme might deal with instances of limited non-compliance. Suppose, for instance, that a prior generation has failed to fulfil certain obligations toward their successors, but not in such a way as to warrant draconian sanctions. In that case, we can conceive of the later generation having the ability to partially resile from their obligations to honour the posthumous interests of their predecessors. To elaborate on this, consider Gosseries's metaphor of the castle which is representative of a political society over time.¹⁶⁵ In keeping with the wider theory of intergenerational mutuality, let us assume that each generation represents a separate group of individuals who own and occupy the castle for a set period of time. Suppose now, that this castle is filled with innumerable projects and undertakings carried out by previous residents – some of these may be trivial, but others will have had enormous significance to the people in question. During their tenure, the cohort can use the castle as they fit, so long as they leave it in as good a condition as they received it – the just savings principle, when applied this example – and show an appropriate degree of respect toward the extant projects of previous inhabitants. In return, they can contribute projects of their own to the castle. Now, for the sake of the argument, let us imagine that one group of occupants arrive to find that the previous inhabitants have left the castle in a state of minor disrepair. In those circumstances, it would be reasonable to assume that the projects of the non-compliant former residents would not be owed the same degree of respect as they might have otherwise had. Notice, however, that this sanctioning mechanism does not require the practices which regulate occupancy in the castle to break down: the new residents are not required to

¹⁶⁵ Axel Gosseries, 'Three Models of Intergenerational Reciprocity' in Gosseries and Meyer, *Intergenerational Justice* p. 132.

expunge every project inherited from their predecessors. In much the same way, we can conceive of limited non-compliance in the intergenerational context being dealt with similarly: later generations can still observe the general practice of honouring the posthumous interests of the predecessors, albeit in a less vigorous and somewhat attenuated manner. Indeed, at the level of a political society we can – and frequently do – choose to impede certain projects inherited from our forebears whilst upholding others. In ideal circumstances, the project being hampered would be closely associated with the act of historical non-compliance. However, even in circumstances where this is not possible, the fact remains that later generations can sanction their ancestors in cases of limited non-compliance. As we will now see, though, the forms of sanction on offer to later generations become increasingly ruinous as non-compliance becomes more widespread.

Widespread non-compliance poses a unique challenge to a theory of intergenerational mutuality for the simple reason that it is not possible to sanction earlier generations without collapsing the reciprocal scheme. To illustrate the nature of the problem, let us return to Gosseries's castle. Now suppose that our cohort arrives to find the castle in a state of absolute dilapidation as a direct result of the previous residents' disregard for future occupants. In these circumstances, the only adequate form of sanction would involve showing a comparable disregard for the projects that the non-compliant residents had left behind. As we have discussed, though, this would spell the end of the practices which, up until that point, had regulated occupancy within the castle. There may, however, be an additional question which is worthy of consideration: would the new cohort abandon the castle, or might they be willing to restore it and start the process over again? If the ability to memorialise their achievements and pass their projects on is as important as I have argued, then they would have strong reasons to prefer the latter option. Indeed, in this respect, there is a clear parallel with our earlier

discussion of the first generation who initiate those practices associated with intergenerational mutuality. Recall, that this initial generational cohort receive no inheritance from their predecessors but have reasons of their own to start the process of intergenerational savings: they have a general interest in bringing about fair intergenerational norms and a more specific one in furthering their posthumous interests. On the basis of that analysis, we can infer that any generation who has to collapse one particular iteration of the scheme to punish compliance would seek to revive it. Indeed, in this respect, it is helpful to consider an analogy with political societies: over time, they reconstitute themselves through processes of comprehensive reform or revolution. In much the same way, we can interpret the process of winding up the reciprocal scheme in response to widespread non-compliance as something which may very well happen within a wider system of intergenerational cooperation. However, so long as the generation who have to administer this sanction are willing to restart the process, then this is unproblematic. So much to say, then, that intergenerational mutuality endures even if specific iterations of the indirect reciprocal scheme come to an end. From this, we can conclude that effective enforcement mechanisms are present with the theory and that, as a consequence, the parties to the original position can be assured that non-compliance will be met with sanctions.

5.5 The First Generation, Non-Compliance, and Reciprocity

At this point, we have addressed the four objections set out in Section 1 and demonstrated the resilience of intergenerational mutuality in the face of such criticism. It would, however, be remiss not to acknowledge that in the process of responding to these lines of argument a further tension has emerged. As we have seen, generational cohorts are able to sanction widespread non-compliance because even if the scheme collapses, they will have reasons to reconstitute it. That, however, jars with the reciprocal norms which undergird the wider theory of

intergenerational mutuality. If, for instance, generations have their own independent reasons to transfer capital to the next generation, then we have to question whether or not the scheme is characterised by reciprocity in the way that has been suggested. Suppose, for instance, that generations are simply engaged in a process of intergenerational saving to further their own interests: although the pattern of transfers might have the appearance of reciprocal interaction, it would be wholly inaccurate to characterise it as such. In this final section, I therefore want to respond to this in two ways: firstly, I will elucidate the motives of the parties by appealing to the idea of strong reciprocity; and secondly, I will explain why, in ordinary circumstances, the logic of intergenerational reciprocity would take precedent over these reasons. In this way, it will reconcile the independent motivations of generations outside of the scheme with indirect reciprocal interaction within it.

In order to meet this challenge, my first line of response will be to suggest that there is strong reciprocal content in the motives of the first generation. In this way, it will be possible to push back against the claim that reciprocal obligations are only possible within the context of an established institutional arrangement. To appreciate how this is possible, it would be helpful to revisit our discussion of strong and weak forms of reciprocity in the previous chapter. Recall that strong reciprocal interaction revolves around the maintenance and reinforcement of social norms: an action can have reciprocal content even if it yields no future benefits. These interactions are ‘reciprocal’ to the extent that they strengthen social norms from which we all benefit. Consequently, if saving on behalf of future generations is considered a ‘fair’ or ‘just’ course of action, then it would have reciprocal content. The payoff, in this context, consists in the reinforcement of those distributive norms which regulate a society’s major social institutions. Consequently, even outside of an established institutional arrangement, the

decision of any first generation to save on behalf of their successors can have reciprocal content.

Those generations who initiate the savings process can, therefore, be doing so on a reciprocal basis. That, however, leaves open the problem of overdetermination within the scheme of reciprocal interaction. To outline, generations may have more than one reason to participate in a savings process and to further the posthumous interests of their forebears: unless the logic of intergenerational reciprocity prevails over these additional motives, then the scheme rests upon shaky ground. What, in other words, is the factor which hands primacy to these reciprocal obligations in the intergenerational context? To answer this, it would be helpful to revisit our prior discussion of the just savings principle and social bases of self-respect. There, it was argued that the just savings principle emerges over its competitors by virtue of its status as a guarantor of the social bases of self-respect. Elsewhere, we have discussed the corresponding relationship between the honouring of posthumous interests and the social bases of self-respect. The reciprocal arrangement between generations is therefore closely associated with the promotion and safeguarding of what Rawls famously considered to be ‘perhaps the most important primary good,’ without which our ability to develop and act upon a conception of the good is greatly reduced.¹⁶⁶ To elaborate, any generation finding itself in the position of having to initiate a savings process does so from a disadvantageous position: their predecessors were either incapable of making an equitable bequest toward them or failed to comply with the demands of justice. Thus, while they may have strong, independent reasons to save for future generations, they do so from a position of relative hardship. By way of contrast, a generation whose predecessors have complied will receive a fair bequest from their predecessors – which, as we have argued, is conducive to the self-respect of the people who comprise that generation.

¹⁶⁶ Rawls, *Theory* p. 348.

Moreover, they will be able to make a full contribution to the scheme by saving for their successors and honouring the posthumous interests of their ancestors. In summary, appreciable benefits can be derived from relating to past and future generations on an equitable basis. For this reason, the reciprocal logic that was outlined in the previous chapter will take precedence over those independent rationales for saving. Under normal circumstances, therefore, the system of intergenerational mutuality will be maintained on an explicitly reciprocal basis.

5.6 Conclusion

To conclude, this chapter has demonstrated that the theory of intergenerational mutuality is able to withstand the four major internal objections to it. In Sections 2 and 3, the incorporation of posthumous interests into the contract method was defended with reference to the social bases of self-respect. Indeed, the relationship between this specific primary good, posthumous interests, and the just savings principle has been one of the major threads running through this chapter. As we have seen, there is a symmetry between the honouring of posthumous interests and the just savings principle which can be explained with reference to the social bases of self-respect: in outline, both can be understood to safeguard and promote them within the intergenerational context. In Section 4, attention turned to the ability of the reciprocal scheme to reconstitute itself in the event of widespread non-compliance. Then, in Section 5 we were able to draw these narrative threads together to further explicate the nature of intergenerational reciprocity. In this regard, we ought to acknowledge the centrality of the social bases of self-respect to the wider theory of intergenerational mutuality: they are the primary good which, on a contractualist analysis, makes the wider system of cooperation between generations both possible and necessary. In the final analysis, then, a deeper understanding of mutuality as it

applies within the intergenerational context allows for a richer appreciation of two of the central concepts within contractualist theory: the social bases of self-respect and reciprocity itself.

Conclusion

In the introductory chapter, I identified two problems which militate against the extension of contractualism to the intergenerational context: the non-identity and mutuality problems, respectively. To that end, this thesis has proceeded in two distinct phases. In Part I, I demonstrated that the non-identity problem can be overcome in a manner that is entirely consistent with the contract method: by distinguishing wrongdoing from harming, we can object to the non-identity cases whilst retaining a person-affecting approach to questions of morality and justice. In addition, I suggested that we could use such a framework to establish our humanitarian obligations toward future generations, on the grounds that they will have rights. I began Part II with a critical assessment of the universal contract approach. In turn, this allowed me to demonstrate the preferability of an account grounded in reciprocity – and, by implication, the centrality of the mutuality problem to the thesis as a whole. To that end, a theory of intergenerational mutuality was expounded in Chapter 4. As against the claim that the intergenerational context is devoid of mutuality, I demonstrated that there is a very real sense in which we are vulnerable to posterity: many of our major life projects and conceptions of the good are dependent on the compliance of future generations in a wider system of intergenerational cooperation. To illustrate how this insight could be incorporated into the contract approach, I suggested that this form of mutually advantageous cooperation had significance from the perspective of stability and the social bases of self-respect: consequently, mutually disinterested parties would be minded to specify an intergenerational savings principle in the original position. Having outlined the theory of intergenerational mutuality, I then proceeded to examine Brian Barry's objection to indirect reciprocity as a basis for obligations at the intergenerational level. In meeting the justificatory challenge, I was able to

give an indication of the explanatory potential of intergenerational mutuality in our deliberations on what is owed to future generations. Finally, I considered objections to intergenerational mutuality and, by responding to them drew attention to two fundamental aspects of the theory: firstly, the centrality of the social bases of self-respect to our understanding of indirect reciprocal interaction between generations; and secondly, the ability of the scheme to reconstitute itself in extreme circumstances.

Intergenerational mutuality is, then, the idea at the heart of this thesis. The claim, in summary, is that past and future generations are reliant on one another's cooperation in order to secure certain essential goods. To that end, I have taken an account of lifetime-transcending interests from the works of Janna Thompson and Samuel Scheffler and married it to the Rawlsian concept of the social bases of self-respect in order to locate it within the contractualist framework. If my analysis is correct, then much theorising on the subject of intergenerational justice rests on a falsehood: that relations between past, present, and future generations can never exhibit or be characterised by mutuality. That, of course, is not to say that the concept of intergenerational mutuality can single-handedly resolve the myriad problems that arise in the intergenerational context. Instead, it would be better to think of it as allowing for a richer and more varied understanding of intergenerational cooperation. To that extent, then, it should be seen as an important and significant step toward a better understanding of our relations at the intergenerational level. I would also want to suggest that there is a distinctly practical element to a concept like intergenerational mutuality, in at least two regards: firstly, by emphasising the mutual dependence that obtains between generations, it might reasonably be hoped that it can motivate us to take a greater interest in the question of justice between generations. Secondly, though, it raises the question of what, if anything, could be done to promote the sense of mutuality between generations. Given the exigent need

to confront issues such as climate change, appeals to our dependence on future generations – and, for that matter, on there being a future at all – may yet prove to be an effective spur to action.

It would, of course, be remiss not to reflect on the limitations of this thesis – and the prospects for further research. I therefore want to attend to three separate issues in the wider intergenerational context. The first of these relates to my decision to focus specifically on distributive justice. As I stated in my introductory chapter, a theory of fair shares is but one part of a much broader theory of relational or social egalitarianism. It would, therefore, be interesting to see how a concept like intergenerational mutuality might be applied in related subject areas. In this respect, the relationship between indirect reciprocity and the more substantive concept of environmental stewardship described by Edward A. Page, along with Dominic Welburn's recent work linking stewardship to a conception of Rawlsian political liberalism would represent a promising point of departure.¹⁶⁷ Secondly, although I stand by my decision to limit the scope of this enquiry to the question of justice and future generations, a comprehensive theory of intergenerational mutuality will need to take account of our obligations toward our predecessors. This could – and, to my mind, should – seek to extend the concept of intergenerational mutuality to the realm of reparative justice. In that way, we could look to extend intergenerational mutuality to non-ideal theory. Finally, by virtue of the fact that issues like global warming are international in scope, it would be interesting to consider how a concept like intergenerational mutuality might be applied at the international-cum-global level. As I suggested in my discussion of the problem of scope in Chapter 4, a theory of intergenerational mutuality could engender obligations based on a

¹⁶⁷ Edward A. Page, 'Fairness on the Day after Tomorrow: Justice, Reciprocity and Global Climate Change', *Political Studies* 55 (2007): 225-42 pp. 232-8; Dominic Welburn, 'Rawlsian Environmental Stewardship and Intergenerational Justice', *Environmental Ethics* 36(4) (2014): 387-404.

number of distributive norms. I am therefore minded to think that it could be incorporated into a number of institutionalist approaches.

The idea of intergenerational mutuality is, then, a one with manifold applications in a number of contexts. Given its explanatory and motivational potential, my hope would be that it might evolve into a broader theory linking theory and practice. At this early juncture, however, that must be read as little more than a speculative aside. For our present purposes, the salient conclusion to draw from this enquiry is that the mutuality problem very possibly rests on a misconception. To the extent that it has proven to be an impediment to theories of intergenerational justice, this may give us cause for cautious optimism.

Bibliography

Anderson, Elizabeth (1999). 'What Is the Point of Equality', *Ethics* 109(2): 287-337.

Archer, David; Eby, Michael; Brovkin, Victor; Ridgwell, Andy; Cao, Long; Mikolajewicz, Uwe;... Tokos, Kathy (2009). Atmospheric lifetime of fossil fuel carbon dioxide', *Annual Review of Earth and Planetary Sciences* 37: 117-34.

Arrhenius, Gustaf (1999). 'Mutual Advantage Contractarianism and Future Generations', *Theoria* 65(1): 25-35

Attas, Daniel (2009). 'A Transgenerational Difference Principle' in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (Oxford: Oxford University Press), 189-218.

Baier, Annette (1981). 'The Rights of Past and Future Persons', in *Responsibilities to Future Generations. Environmental Ethics*, Ernest Partridge (ed.) (New York, NY.: Prometheus Books, 1981), 171-83.

Barry, Brian (1989). *A Treatise on Social Justice, Vol. I: Theories of Justice* (Berkeley, CA.: University of California Press).

— (1991a). 'Justice as Reciprocity', in *Liberty and Justice: Essays in Political Theory 2* (Oxford: Clarendon Press), 211-41.

— (1991b). 'Justice Between Generations', in *Liberty and Justice: Essays in Political Theory* 2 (Oxford: Clarendon Press), 242-58.

Beitz, Charles R. (1983). 'Cosmopolitan Ideals and National Sentiment,' *Journal of Philosophy* 80(10): 591-600.

Benatar, David (2006). *Better Never to Have Been: The Harm of Coming Into Existence* (Oxford: Clarendon Press).

Blake, Michael and Smith, Patrick (2015). 'International Distributive Justice', *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), ed. Edward N. Zalta, URL:
<https://plato.stanford.edu/archives/spr2015/entries/international-justice>

Brandstedt, Eric (2015). 'The Circumstances of Intergenerational Justice', *Moral Philosophy and Politics* 2(1): 33-56.

— (2017). 'The savings problem in the original position: assessing and revising a model', *Canadian Journal of Philosophy* 47(2) (2017): 269-289.

Brooks, Roy L. (2004). 'Getting Reparations for Slavery Right – A Response to Posner and Vermeule', *Notre Dame Law Review* 80 (1): 251-88.

Burke, Edmund (1968). *Reflections on the Revolution in France* (London: Penguin).

Caney, Simon (2010). 'Climate Change, Human Rights, and Moral Thresholds', in *Climate Ethics: Essential Readings* eds. Stephen M. Gardiner, Simon Caney, Dale Jamieson, Henry Shue (New York, NY.: Oxford University Press) pp. 122-45.

Darwall, Stephen (2003). 'Introduction' to *Contractarianism/Contractualism* (Oxford: Blackwell), 1-8.

Dworkin, Ronald (1978). *Taking Rights Seriously* (Cambridge, MA.: Harvard University Press).

English, Jane (1977). 'Justice Between Generations', *Philosophical Studies* 31: 91-104

Finneron-Burns, Elizabeth (2016). 'Contractualism and the Non-Identity Problem', *Ethical Theory and Moral Practice* 19(5): 1151-63.

— (2017). 'The Intergenerational Original Position', *Social Theory and Practice* 43(4): 805-23.

Gardiner, Stephen M. (2009). 'A Contract on Future Generations', in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (Oxford: Oxford University Press), 77-118.

— (2011). *A Perfect Moral Storm: The Ethical Tragedy of Climate Change* (New York, NY.: Oxford University Press).

Gauthier, David (1963). *Practical Reasoning: The Structure and Foundations of Prudential and Moral Arguments and Their Exemplification in Discourse* (Oxford: Clarendon Press).

— (1986). *Morals by Agreement* (New York, NY: Oxford University Press).

Gintis, Herbert; Bowles, Samuel; Boyd, Robert; and Fehr, Ernst (2005). 'Moral Sentiments and Material Interests: Origins, Evidence, and Consequences' in Herbert Gintis, Samuel Bowles, Robert Boyd, and Ernst Fehr (eds.) *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life* (Cambridge, MA.: The MIT Press).

Gosseries, Axel (2009). 'Three Models of Intergenerational Reciprocity' in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (New York, NY.: Oxford University Press), 119-146.

Heath, Joseph (1997). 'Intergenerational Cooperation and Distributive Justice', *Canadian Journal of Philosophy* 27(3): 361-76.

— (2013). 'The Structure of Intergenerational Cooperation', *Philosophy and Public Affairs* 41(1): 31-66.

Heyd, David (1992). *Genethics: Moral Issues in the Creation of People* (Berkeley, CA.: University of California Press).

— (2009). ‘A Value or an Obligation? Rawls on Justice to Future Generations’, in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (Oxford: Oxford University Press), 167-88.

Hume, David (1983). *An Enquiry Concerning the Principles of Morals* ed. J.B. Schneewind (Indianapolis, IN.: Hackett Publishing Company).

Jefferson, Thomas (1984). ‘Letter to James Madison, Paris, 6th September, 1789’ in *Thomas Jefferson: Writings* ed. Merrill D. Peterson (New York, NY.: Literary Classics).

Karnein, Anja (2015). ‘Climate Change and Justice Between Nonoverlapping Generations’, *Global Justice: Theory Practice Rhetoric* 8(2): 43-65.

Katz, Corey (2017). ‘Contractualism, Person-Affecting Wrongness, and the Non-Identity Problem’, *Ethical Theory and Moral Practice* 21(1): 103-19.

Korsgaard, Christine M. (1989). ‘Personal Identity and the Unity of Agency: A Kantian Response to Parfit’, *Philosophy and Public Affairs* 18(2) (1989): 101-32.

Kumar, Rahul (2003). ‘Who Can Be Wronged?’, *Philosophy and Public Affairs* 31(2): 99-118.

— (2009). ‘Wronging Future People: A Contractualist Proposal’, in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (New York, NY.: Oxford University Press), 251-272.

Mathis, Klaus (2009). 'Future Generations in John Rawls' Theory of Justice', *Archives of Law and Social Philosophy* 95(1): 49-61.

Miller, David (1995). *On Nationalism* (Oxford: Clarendon Press).

Mills, Charles (2007). 'Contract of Breach: Repairing the Racial Contract' in Carol Pateman and Charles Mills, *Contract and Domination* (Cambridge: Polity Press).

Okin, Susan Moller (1989). *Justice, Gender, and the Family* (New York, NY: Basic Books).

Page, Edward A. (2007). 'Fairness on the Day after Tomorrow: Justice, Reciprocity and Global Climate Change', *Political Studies* 55: 225-42.

Paine, Thomas (1995). 'The Rights of Man' in *Rights of Man, Common Sense, and Other Political Writings* ed. Mark Phillp (New York, NY.: Oxford University Press).

Parfit, Derek (1986). 'Comments', *Ethics* 96(4): 832-72.

— (1987). *Reasons and Persons* (Oxford: Clarendon Press).

— (2011). *On What Matters, Vol. 2* (Oxford: Oxford University Press).

Rawls, John (1971). *A Theory of Justice* (Cambridge, MA.: Harvard University Press)

— (1993). *Political Liberalism* (New York, NY.: Columbia University Press).

— (1999a). *The Law of Peoples* (Cambridge, MA.: Harvard University Press).

— (1999b). *A Theory of Justice* Revised Edition (Cambridge, MA: Belknap).

— (2001). *Justice as Fairness: A Restatement* ed. Erin Kelly (Cambridge, MA: Harvard University Press).

Reiman, Jeffrey (2007). 'Being Fair to Future People': The Non-Identity Problem in the Original Position', *Philosophy and Public Affairs* 35(1): 71–92.

Reglitz, Merten (2016). 'The Practice-Independence of Intergenerational Justice', *Utilitas* 28(4): 415-40.

Ron, Amit (2006). 'Rawls as a critical theorist: Reflective equilibrium after the 'critical turn'', *Philosophy and Social Criticism* 32(2): 173-91.

Sangiovanni, Andrea (2007). 'Global Justice, Reciprocity, and the State,' *Philosophy and Public Affairs* 35: 3-39.

Scanlon, T.M. (1998). *What We Owe to Each Other* (Cambridge, MA.: Harvard University Press).

Scheffler, Samuel (2003). 'What is Egalitarianism?', *Philosophy and Public Affairs* 31(1): 5-39.

— (2013). *Death and the Afterlife* ed. Niko Kolodny (New York, NY.: Oxford University Press).

— (2018). *Why Worry About Future Generations?* (Oxford: Oxford University Press).

Shiffrin, Seana (1999). 'Wrongful Life, Procreative Responsibility, and the Significance of Harm', *Legal Theory* 5: 117-48

Smith, Steven R. (2001). The Social Construction of Talent: A Defence of Justice as Reciprocity', *Journal of Political Philosophy* 9(1): 19-37.

Tremmel, Joerg Chet (2009). *A Theory of Intergenerational Justice* (London: Earthscan).

Thoma, Johanna (2015). 'Bargaining and the Impartiality of the Social Contract' *Philosophical Studies* 172: 3335-3355.

Thompson, Janna (2009a). 'Identity and Obligation in a Transgenerational Polity', in *Intergenerational Justice* eds. Axel Gosseries and Lukas H. Meyer (New York, NY.: Oxford University Press), 25-49.

— (2009b). *Intergenerational Justice: Rights and Responsibilities in an Intergenerational Polity* (New York, NY: Routledge).

Vernon, Richard (2016). *Justice Back and Forth: Duties to the Past and Future* (Toronto: Toronto University Press).

Visser 't Hooft, H.P. (1999). *Justice to Future Generations and the Environment* (Boston, MA.: Kluwer Academic Publishers).

Warren, Paul (1994). 'Self-Ownership, Reciprocity, and Exploitation, or Why Marxists Shouldn't Be Afraid of Robert Nozick', *Canadian Journal of Philosophy* 24(1): 33-56.

Weinberg, Rivka (2008). 'Identifying and Dissolving the Non-Identity Problem', *Philosophical Studies* 137: 3-18.

Welburn, Dominic (2014). 'Rawlsian Environmental Stewardship and Intergenerational Justice', *Environmental Ethics* 36(4): 387-404.

Wolf, Susan (2013). 'The Significance of Doomsday', in *Death and the Afterlife* ed. Niko Kolodny (New York, NY.: Oxford University Press), 113-30.

Willetts, David (2010). *The Pinch: How the baby boomers took their children's future – and why they should give it back* (London: Atlantic Books).

Woodward, James (1986). 'The Non-Identity Problem', *Ethics* 96(4): 804-31.